Г.Е. ШИЛОВ

О С Н О В Ы

М А Т Е М А Т И Ч Е С К О Г О А Н А Л И З А

(Методическое пособие)

Выпуск 3 .

§ 3.I. Общее определение высших производных.

3.II. а. Как мы уже хорошо знаем, для функции $f(x): G \subset X \to Y$ линейный оператор $f'(a): X \to Y$ определяется из равенства

$$f(a+h) - f(a) = f'(a)h + o(h), \quad h \in X.$$

Предположим, что $f'(x)$ существует в каждой точке $x \in G$. Тогда $f'(x)$ есть функция, ставящая в соответствие каждой точке $x \in G$ оператор $X \to Y$. Пространство $L(X, Y)$ всех линейных непрерывных операторов, действующих из X в Y, мы обозначим здесь через Y_1 (а само Y – через Y_0). Таким образом, $f(x)$ есть Y_0-значная функция от x, а $f'(x)$ есть Y_1-значная функция от x.

Определим теперь оператор $f''(a)$ из равенства (если оно возможно)

$$f'(a+h) - f'(a) = f''(a)h + o(h), \quad h \in X.$$

Тогда $f''(a)$ будет линейным непрерывным оператором, действующим из X в Y_1. Пространство $L(X, Y_1)$ всех таких операторов обозначим через Y_2. Если $f''(x)$ существует для любого $x \in G$, то она будет Y_2-значной функцией от x. Продолжая далее, мы приходим к следующему общему определению.

О п р е д е л е н и е. Пусть $Y_p = L(X, Y_{p-1})$ $(p = 1, 2, \ldots)$; линейный оператор $f^{(p)}(a): X \to Y_{p-1}$ называется p-й производной от функции $f(x)$ $(G \subset X \to Y_0)$ в точке $x = a$, если справедливо представление

$$f^{(p-1)}(a+h) - f^{(p-1)}(a) = f^{(p)}(a)h + o(h) \quad (h \in X) \qquad (1)$$

Все слагаемые в обеих частях равенства здесь принадлежат к пространству Y_{p-1}. Таким образом, $f^{(p)}(a) \in Y_p$.

Коротко можно написать

$$f^{(p)}(a) = \left[f^{(p-1)}(x)\right]'\Big|_{x=a} \qquad (2)$$

и это приводит определение p-й производной в конечном счете к определению первой производной.

б. Функция $y = f(x)$ $(G \subset X \to Y)$, имеющая при $x = a \in G$ производные до порядка p включительно, называется p раз дифференцируемой, короче, p-дифференцируемой в точке a; если она имеет производные до порядка p вклю-

чительно в каждой точке области $G \subset X$, то она называется p раз дифференцируемой (p-дифференцируемой) в области G. Функция $f(x)$, имеющая при $x = a$ (при каждом $x \in G$) производные любого порядка $p = 1, 2, \ldots$, называется бесконечно дифференцируемой в точке a (в области G).

Функция, p-дифференцируемая в области G и имеющая непрерывные в G производные $f'(x), \ldots, f^{(p)}(x)$, называется функцией класса $C^p(G)$ или короче C^p-функцией. Функция, имеющая непрерывные производные любого порядка в области G, называется C^∞-функцией.

в. Если функция $f(x)$ имеет в точке a производные до порядка k, а ее k-я производная имеет в этой же точке производные до порядка m, то функция $f(x)$ при $x = a$ имеет производные до порядка $k+m$ и

$$f^{(k+m)}(a) = \left[f^{(k)}(x) \right]^{(m)} \Big|_{x=a} \qquad (3)$$

Действительно, если $m = 1$, то сказанное есть просто определение функции, дифференцируемой $k+1$ раз; в общем случае сформулированный результат легко доказывается по индукции.

Обращение полученного предложения очевидно: если функция $f(x)$ дифференцируема $k+m$ раз при $x = a$, то $f^{(k)}(x)$ дифференцируема m раз и также имеет место формула (3).

3.12. Если $X = R_1$, то всякий линейный оператор $X \to Y$ отождествляется естественно с элементом пространства Y, так что мы имеем $Y \equiv Y_0 = Y_1 = Y_2 = \ldots$ (I.24б). Поэтому для функции $f(x)$ вещественного переменного x со значениями в пространстве Y все производные $f'(x), f''(x), \ldots$ также принимают свои значения в пространстве Y.

3.13 а. Рассмотрим случай $X = R_n$, $Y = R_1$. Здесь $Y_1 = L(R_n, R_1) = R_n$, $Y_2 = L(R_n, R_n) = R_n^2$ и т.д. Оператор $f'(x)$ может быть задан n составляющими:

$$\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n};$$

этот оператор действует на вектор смещения $h = dx = (dx_1, \ldots, dx_n)$ по формуле I.22 (3):

$$f'(x)\,dx = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i}\, dx_i.$$

Чтобы найти $f''(x)$, мы должны из вектора приращения $f'(x)$

$$f'(x+h)-f'(x)=\left\|\begin{array}{c}\frac{\partial f(x+h)}{\partial x_1}-\frac{\partial f(x)}{\partial x_1}\\ \cdots\cdots\cdots\\ \frac{\partial f(x+h)}{\partial x_n}-\frac{\partial f(x)}{\partial x_n}\end{array}\right\| \qquad (1)$$

выделить главную линейную часть; это равносильно выделению главной линейной части из каждой строки вектора $f'(x+h)-f'(x)$, что приводит к матрице

$$\left\|\begin{array}{ccc}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots \frac{\partial^2 f}{\partial x_n \partial x_1}\\ \cdots & \cdots & \cdots\\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots \frac{\partial^2 f}{\partial x_n^2}\end{array}\right\| \qquad (2)$$

Этой матрицей задается оператор $f''(x): X \to Y_1 = R_n$, действующий на вектор $h=(h_1,\dots,h_n)\in R_n = X$ по формуле

$$f''(x)h=\left\|\begin{array}{c}\frac{\partial^2 f}{\partial x_1^2}h_1+\frac{\partial^2 f}{\partial x_2 \partial x_1}h_2+\cdots+\frac{\partial^2 f}{\partial x_n \partial x_1}h_n\\ \cdots\cdots\cdots\\ \frac{\partial^2 f}{\partial x_1 \partial x_n}h_1+\frac{\partial^2 f}{\partial x_2 \partial x_n}h_2+\cdots+\frac{\partial^2 f}{\partial x_n^2}h_n\end{array}\right\| \in Y_1$$

В свою очередь оператор $f''(x)h \in Y_1 = L(R_n, R_1)$ действует на вектор $k=(k_1,\dots,k_n)\in R_n = X$ по формуле

$$f''(x)hk=\left(\frac{\partial^2 f}{\partial x_1^2}h_1+\cdots+\frac{\partial^2 f}{\partial x_n \partial x_1}h_n\right)k_1+\cdots+$$

$$+\left(\frac{\partial^2 f}{\partial x_1 \partial x_n}h_1+\cdots+\frac{\partial^2 f}{\partial x_n^2}h_n\right)k_n=\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}h_i k_j \qquad (3)$$

Как мы знаем, существование $f'(x)$ влечет существование частных производных $\frac{\partial f(x)}{\partial x_1},\dots,\frac{\partial f(x)}{\partial x_n}$. Аналогично, существование $f''(x)$ влечет существование всех частных производных второго

порядка $\frac{\partial^2 f}{\partial x_i \partial x_j}$ $(i,j=1,\dots,n)$. Продолжая таким же образом, мы видим, что существование оператора $f^{(p)}(x)$ означает дифференцируемость всех частных производных порядка $p-1$ у функции $f(x)$, в частности, влечет существование всех частных производных порядка p.

6. Эти результаты не имеют прямого обращения: из наличия частных производных $\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n}$ еще не следует существования $f'(x)$ (I.47); тем более, из наличия вторых частных производных $\frac{\partial^2 f}{\partial x_i \partial x_j}$ не следует существования $f''(x)$, и т.д. Однако, если первые производные существуют и непрерывны, то существование (и непрерывность) $f'(x)$ уже обеспечивается теоремой I.47. Аналогично, если производные $\frac{\partial^2 f}{\partial x_i \partial x_j}$ существуют и непрерывны, то существуют и непрерывны первые производные $\frac{\partial f}{\partial x_j}$; следовательно, существует и непрерывна $f'(x)=\left\{\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right\}$; далее, по тому же предположению существуют и непрерывны производные $\frac{\partial}{\partial x_j} f'(x)=\left\{\frac{\partial^2 f}{\partial x_1 \partial x_j}, \dots, \frac{\partial^2 f}{\partial x_n \partial x_j}\right\}$, откуда, в силу той же теоремы I.47, функция $f'(x)$ является дифференцируемой, так что существует – и, очевидно, непрерывна – $f''(x)$; таким образом существование и непрерывность $f'(x)$ в области G равносильны существованию и непрерывности всех частных производных второго порядка $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ в области G.

Эти рассуждения легко обобщаются и на высшие производные.

Существование и непрерывность оператора $f^{(p)}(x)$ в области G равносильны существованию и непрерывности всех частных производных порядка p от функции $f(x)$ в области G.

3.I4. **Полилинейные и степенные формы.**

а. С помощью линейного оператора $A_1 : X \to Y = Y_0$ и вектора $h_1 \in X$ можно образовать вектор $A_1 h_1 \in Y_0$. С помощью линейного оператора $A_2 : X \to Y_1 = L(X, Y_0)$ и вектора $h_2 \in X$ образуется оператор $A_2 h_2 \in L(X, Y_0)$, и, далее, с помощью вектора $h_1 \in X$ можно образовать выражение $(A_2 h_2) h_1 \in Y_0$. Это выражение есть Y_0-значная билинейная форма от векторов

h_1 и h_2. Продолжая аналогично, с помощью линейного оператора $A_p : X \to Y_{p-1} = L(X, Y_{p-2})$ и векторов $h_1, \dots, h_p$ пространства X можно образовать выражение

$$A_p h_p h_{p-1} \dots h_1 = (\dots((A_p h_p) h_{p-1}) \dots h_1) \in Y_0,$$

которое представляет собой Y_0-значную p-линейную форму от векторов $h_1, \dots, h_p$.

6. Если $A_p : X \to Y_{p-1}$ — ограниченный линейный оператор, то, в соответствии с общим определением нормы оператора (012.6I6), имеет место оценка

$$|A_p h_p h_{p-1} \dots h_1| \leq \|A_p h_p \dots h_2\| \cdot |h_1| \leq$$
$$\leq \|A_p\| \cdot |h_p| \dots |h_1| ;$$

следовательно,

$$\sup_{|h_1| \leq 1, \dots, |h_p| \leq 1} |A_p h_p \dots h_1| \leq \|A_p\|. \qquad (1)$$

С другой стороны, пусть θ, $0 < \theta < 1$, — фиксированное число; если дан оператор $A_p : X \to Y_{p-1}$, то существует вектор h_p, $|h_p| = 1$, для которого $\|A_p h_p\| \geq \theta \|A_p\|$; далее имеется вектор h_{p-1}, $|h_{p-1}| = 1$, для которого $\|A_p h_p h_{p-1}\| \geq \theta \|A_p h_p\| \geq \theta^2 \cdot \|A_p\|$; продолжая это построение, мы придем к последовательности нормированных векторов $h_p, \dots, h_1$, для которых выполняется неравенство

$$|A_p h_p h_{p-1} \dots h_1| \geq \theta^p \cdot \|A_p\|.$$

Отсюда

$$\sup_{|h_1| \leq 1, \dots, |h_p| \leq 1} |A_p h_p \dots h_1| \geq \theta^p \cdot \|A_p\|,$$

и так как $\theta \in (0,1)$ произвольно, то

$$\sup_{|h_1| \leq 1, \dots, |h_p| \leq 1} |A_p h_p \dots h_1| \geq \|A_p\|. \qquad (2)$$

Сравнивая (I) и (2), получаем, что

$$\|A_p\| = \sup_{|h_1| \leq 1, \dots, |h_p| \leq 1} |A_p h_p \dots h_1|.$$

8. Если положить $h_1 = \dots = h_p = h$, то мы получим степенную форму

$$A_p h \dots h.$$

Форма $A_p h_p \dots h_1$ называется симметричной p-формой, если ее значение не изменяется при любой перестановке векторов $h_1, \dots, h_p$.

Значения полилинейной симметричной формы $A_p h_p \dots h_1$ можно найти, зная значения соответствующей степенной формы $A_p h \dots h$. Вот простой, хотя, быть может, не самый экономный способ такого вычисления. Имея векторы $h_1, \dots, h_p$, рассмотрим значения степенной формы $A_p h \dots h$ на векторе $h = h_1 + \dots + h_p = x$:

$$(I) = A_p (h_1 + \dots + h_p) \dots (h_1 + \dots + h_p) =$$

$$= \sum_{(k)} C_{k_1 \dots k_p} A_p \underbrace{h_p \dots h_p}_{k_p \text{ раз}} \cdots \underbrace{h_1 \dots h_1}_{k_1 \text{ раз}},$$

где $k_1 + \dots + k_p = p$, (k) означает комплекс из p чисел $k_1 \geqslant 0, \dots, k_p \geqslant 0$, а коэффициенты $C_{k_1 \dots k_p}$, образовавшиеся после упорядочения аргументов (допустимого вследствие предположенной симметрии формы) и приведения подобных членов, – целые положительные числа. Заменим в форме (I) аргумент h_1 на $2h_1$; тогда справа в каждом слагаемом формы (I) появится множитель 2^{k_1}, принимающий наибольшее значение 2^p на члене $A_p h_1 \dots h_1$.

Эту форму обозначим через (II). Разность (III) $= 2^p$(I) – (II) уже не содержит члена, содержащего p раз аргумент h_1; остальные же члены по-прежнему имеют целые положительные коэффициенты. Заменяя в форме (III) снова h_1 на $2h_1$, получаем форму (IV), в коэффициентах которой по сравнению с (III) появляются множители – степени двойки, причем наибольший из них есть 2^{p-1}. Тогда разность (V) $= 2^{p-1}$(III) – (IV) не содержит слагаемых, содержащих $p-1$ раз аргумент h_1. Продолжая так далее, после $p-1$ аналогичных операций мы получим форму (VI), в которой слагаемые, имея целые положительные коэффициенты, содержат аргумент h_1 не более, чем по одному разу. Далее таким же образом в форме (VI) можно освободиться от слагаемых, содержащих более чем по одному разу аргументы $h_2, \dots, h_p$. В итоге получается форма (VII), слагаемые которой, имея целые положительные коэффициенты, содержат аргументы $h_1, \dots, h_p$ не более чем по одному разу. А эти

значит, что форма (VII) состоит лишь из одного слагаемого:

$$(VII) = c_p A_p h_p \dots h_1,$$

где c_p есть целое положительное число. Но по построению форма (VII) есть линейная комбинация значений степенной формы $A_p h^{(i)} \dots h^{(i)}$ при некоторых специально выбранных векторах $h^{(i)}$ (именно $h^{(1)} = h_1 + \dots + h_p$, $h^{(2)} = 2h_1 + \dots + h_p$, $h^{(3)} = 2^2 h_1 + \dots + h_p$ и т.д.)

Этим и доказано утверждение.

г. С л е д с т в и е. Для симметричной p-формы $A_p h_p \dots h_1$ из условия $A_p h \dots h = 0$ (для всех $h \in X$) вытекает, что $A_p = 0$.

Действительно, из $A_p h \dots h = 0$ и в следует, что $A_p h_p \dots h_1 = 0$ для любых $h_1, \dots, h_p$ из X ; в точности по б имеем $\| A_p \| = 0$.

д. С л е д с т в и е. Для каждого p существует такая постоянная $c_p > 0$, что для любой p-формы

$$\| A_p \| = \sup_{|h_1| \le 1, \dots, |h_p| \le 1} |A_p h_p \dots h_1| \le c_p \cdot \sup_{|h| \le 1} |A h \dots h|$$

Действительно, согласно в любая p-форма $A_p h \dots h$ представляется в виде вполне определенной линейной комбинации степенных форм $A_p h^{(i)} \dots h^{(i)}$, причем векторы $h^{(i)}$ принадлежат фиксированному шару, откуда и вытекает оценка

$$|A_p h_p \dots h_1| \le c_p \cdot \sup_{|h| \le 1} |A_p h \dots h|.$$

Если внимательно проследить за доказательством в , то можно оценить постоянную c_p. Число слагаемых построенной линейной комбинации не превосходит p^{p^2}, векторы $h^{(i)}$ по норме не превосходят $2^p \cdot p$, и коэффициенты не превосходят 2^{p^3}, откуда

$$c_p \le 2^{p^3} \cdot 2^p \cdot p^{p^3}$$

е. Следствие г можно усилить следующим образом:

Пусть для симметричной p-формы $A_p h_p \dots h_1$ известно, что при $h \to 0$

$$A_p h \dots h = o(|h|^p),$$

- иначе говоря, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что при $|h| < \delta$ выполняется неравенство

$$|A_p h \dots h| \le \varepsilon |h|^p. \qquad (3)$$

Тогда $A_p = 0$.

Действительно, если $A_p \neq 0$, то по г имеется вектор $h_0 \in X$, для которого $A_p h_0 \dots h_0 = \ell \neq 0$. Тогда на луче $h = t h_0$ $(0 < t < \infty)$ имеем

$$A_p h \dots h = t^p A_p h_0 \dots h_0 = t^p \cdot \ell = |h|^p \cdot \frac{\ell}{|h_0|^p},$$

что противоречит (3); поэтому в действительности $A_p = 0$.

3.15. Дифференциалы высших порядков.

а. Пусть функция $y = f(x)$ $(G \subset X \to Y)$ p раз дифференцируема при $x = a$. С помощью оператора $f'(a) : X \to Y$ и вектора $h \in X$ можно образовать линейную форму

$$df(a) = f'(a)\, h.$$

Это - первый дифференциал функции $f(x)$ при $x = a$, отвечающий смещению h (I.23). Далее, с помощью оператора $f''(a) : X \to Y$ и векторов h_1 и h_2 можно образовать билинейную форму $f''(a)\, h_2 h_1$. Соответствующая ей квадратичная форма

$$f''(a)\, h h = d^2 f(a)$$

называется вторым дифференциалом функции $f(x)$ при $x = a$, отвечающим смещению h. Таким образом мы построим все дифференциалы вплоть до дифференциала p-го порядка

$$d^p f(a) = f^{(p)}(a)\, h \dots h,$$

получающегося из p-линейной формы $f^{(p)}(a)\, h_p \dots h_1$ при $h_1 = \dots = h_p = h$.

б. Найдем выражения этих дифференциалов для функции $y = f(x) : R_n \to R_1$. В этом случае, полагая $h = (dx_1^{(1)}, \dots, dx_n^{(1)})$, мы имеем

$$df(a) = f'(a)\, h_1 = \sum_{i=1}^{n} \frac{\partial f(a)}{\partial x_i}\, dx_i^{(1)}.$$

Пусть теперь $h_2 = (dx_1^{(2)}, \dots, dx_n^{(2)})$. Тогда по 3.13 (2)

$$f''(a)\, h_2 h_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} dx_j^{(2)} dx_i^{(1)},$$

что при $h_1 = h_2 = h = (dx_1, \dots, dx_n)$ приводит к выражению для второго дифференциала функции $f(x)$ в точке a:

$$d^2 f(a) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} dx_i dx_j,$$

которое хорошо известно из элементарного анализа. Аналогично, для p-го дифференциала функции $f(x)$ при $x = a$ получаем совпадение с ее p-м дифференциалом, определенным в классическом анализе.

3.16. Ч а с т н ы е п р о и з в о д н ы е в ы с ш и х п о р я д к о в и ч а с т н ы е д и ф ф е р е н ц и а л ы.

Пусть пространство X представлено в виде прямой суммы q замкнутых подпространств $X_1, \dots, X_q$. Соответственно вектор $x \in X$ представляется в виде $x = x_1 + \dots + x_q$, $x_i \in X_i$, $i = 1, \dots, q$. Пусть функция $f(x): G \subset X \to Y$ дифференцируема p раз в области G.

Оператор $f'(a)$ действует, как мы знаем, из пространства X в пространство Y; его сужение на подпространство X_i совпадает с частной производной $\frac{\partial f(x)}{\partial x_i}$ функции $f(x)$ по подпространству X_i (I.46). Этот оператор, по самому смыслу, применяется к векторам $h_i \in X_i$. Функция $\frac{\partial f(x)}{\partial x_i}$ определена в области G и вместе с функцией $f'(x)$ допускает дальнейшее дифференцирование; ее частная производная по подпространству X_j обозначается через $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$. Этот оператор, естественно, применяется к векторам $h_j \in X_j$. Продолжая далее, мы получаем возможность определить частные производные вида $\frac{\partial^p f(x)}{\partial x_{i_1} \dots \partial x_{i_p}}$. Для этих операторов имеют смысл выражения $\frac{\partial^p f(x)}{\partial x_{i_p} \dots \partial x_{i_1}} h_{i_1} \dots h_{i_p}$, $h_{i_1} \in X_{i_1}, \dots, h_{i_p} \in X_{i_p}$, которые называются частными дифференциалами функции $f(x)$ по подпространствам $X_{i_1}, \dots, X_{i_p}$, соответствующими смещениям $h_{i_1}, \dots, h_{i_p}$.

Для функций конечного числа вещественных переменных ($X = R_n$) эти определения обобщают классические определения высших частных производных и дифференциалов.

§ 3.2. Свойства высших производных.

3.21. а. Пусть линейный оператор A отображает пространство X в пространство Y_{p-1} (3.14), так что можно рассмотреть полилинейную форму

$$A x_p \dots x_1, \quad x_1, \dots, x_p \in X, \tag{1}$$

и соответствующую степенную форму $A x \dots x$, $x \in X$.

Теорема. Если форма (1) симметрична, то функция $f(x) = A x \dots x$ бесконечно дифференцируема и

$$\left.\begin{aligned} f'(x) &= p\, A \underbrace{x \dots x}_{p-1 \text{ раз}} \in Y_1 \\ &\dots\dots\dots\dots\dots\dots \\ f^{(k)}(x) &= p(p-1)\dots(p-k+1)\, A \underbrace{x \dots x}_{p-k \text{ раз}} \in Y_k, \\ &\dots\dots\dots\dots\dots\dots \\ f^{(p)}(x) &= p!\, A \in Y_p, \\ f^{(q)}(x) &= 0 \quad (q > p) \end{aligned}\right\} \tag{2}$$

Доказательство. По свойству полилинейности формы $A x_p \dots x_1$ имеем

$$f(x+h) - f(x) = A(x+h)\dots(x+h) - A x \dots x =$$
$$= A h x \dots x + A x h \dots x + \dots + A x \dots x h + o(h)$$

и в силу симметрии формы $A x_p \dots x_1$

$$f(x+h) - f(x) = p\, A \underbrace{x \dots x}_{p-1 \text{ раз}} h + o(h);$$

таким образом,

$$f'(x) = p\, A \underbrace{x \dots x}_{p-1 \text{ раз}}.$$

Далее можно действовать по индукции, предполагая, что для формы $A \underbrace{x \dots x}_{p-1 \text{ раз}}$ формулы (2) справедливы, и учитывая, что для $p = 1$ результат был доказан еще в I.3I а–б.

Дальнейшие утверждения б, в и г для $p = 1$ были доказаны соответственно в I.32 а, б и в, и доказательство их для любого $p = 1, 2, \ldots$ легко проводятся по индукции.

б. Если функции $f(x) : V \subset X \to Y$ и $g(x) : V \subset X \to Y$ дифференцируемы p раз при $x = a \in V$, то и $s(x) = f(x) + g(x)$ дифференцируема p раз при $x = a$ и

$$s^{(p)}(a) = f^{(p)}(a) + g^{(p)}(a);$$

иными словами, для любого $h \in X$

$$s^{(p)}(a)\underbrace{h \ldots h}_{p\ раз} = f^{(p)}(a)\underbrace{h \ldots h}_{p\ раз} + g^{(p)}(a)\underbrace{h \ldots h}_{p\ раз}.$$

в. Если функция $y(x) : V \to Y$ дифференцируема p раз при $x = a$ и A – линейный непрерывный оператор, действующий из пространства Y в пространство Z, то функция $z(x) = Ay(x)$ также p раз дифференцируема при $x = a$ и $z^{(p)}(a) = A y^{(p)}(a)$, или, что то же, для любого $x \in X$

$$z^{(p)}(a)\underbrace{h \ldots h}_{p\ раз} = A y^{(p)}(a)\underbrace{h \ldots h}_{p\ раз}$$

г. Пусть Y есть нормированная прямая сумма подпространств $Y_{(1)}, \ldots, Y_{(n)}$, так что любая функция $y(x) : V \to Y$ обладает составляющими $y_{(1)}(x) : V \to Y_{(1)}, \ldots, y_{(n)}(x) : V \to Y_{(n)}$.

Тогда из p-кратной дифференцируемости функции $y(x)$ при $x = a$ следует p-кратная дифференцируемость при $x = a$ и каждой составляющей $y_j(x)$ $(j = 1, \ldots, n)$; при этом

$$y^{(p)}(x) = \{y_{(1)}^{(p)}(x), \ldots, y_{(n)}^{(p)}(x)\},$$

где $\{\}$ означают набор составляющих функции $y^{(p)}(x) : V \to Y_p$ (3.II а) в естественном разложении пространства Y_p в прямую сумму (I.I4е)

$$Y_p = Y_{p(1)} + \ldots + Y_{p(n)}.$$

Обратно, из p-кратной дифференцируемости при $x = a$ всех составляющих $y_{(1)}(x), \ldots, y_{(n)}(x)$ вытекает p-кратная дифференцируемость при $x = a$ и самой функции $y(x)$.

3.22. **Симметрия второй производной.**

а. Пусть $y = f(x)$ $(G \subset X \to Y)$ – дважды дифференцируемая функция. Составим билинейную форму $f''(a)hk$ от вы-

торов h и k пространства X. Покажем, что эта билинейная форма симметрична, т.е. для любых векторов h и k имеет место равенство

$$f''(a)\,hk = f''(a)\,kh. \qquad (1)$$

Для доказательства рассмотрим выражение

$$w = f(a+h+k) - f(a+h) - f(a+k) + f(a).$$

Его можно рассматривать как приращение функции

$$\Phi(x) = f(x+k) - f(x)$$

при изменении x от a до $a+h$. По теореме о конечном приращении I.4I (8) находим

$$|\Phi(a+h) - \Phi(a) - \Phi'(a)h| \le \sup_{0<\theta<1} |\Phi'(a+\theta h) - \Phi'(a)|\,|h| \qquad (2)$$

Для заданного $\varepsilon > 0$ найдем $\delta > 0$ такое, чтобы из неравенства $|h| \le \delta$ следовало

$$|f'(a+h) - f'(a) - f''(a)h| \le \varepsilon |h| \le \varepsilon\delta$$

В дальнейших формулах будем брать $|h| \le \frac{\delta}{2}$, $|k| \le \frac{\delta}{2}$ и через $\varepsilon_1, \varepsilon_2, \ldots$ обозначать величины (векторы, операторы) с нормой, меньшей $\varepsilon\delta$.

Мы имеем $\Phi'(x) = f'(x+k) - f'(x)$, так что

$$\Phi'(a+\theta h) = f'(a+k+\theta h) - f'(a+\theta h) =$$
$$= [f'(a+k+\theta h) - f'(a)] - [f'(a+\theta h) - f'(a)] =$$
$$= [f''(a)(k+\theta h) + \varepsilon_1] - [f''(a)\theta h + \varepsilon_2] = f''(a)\,k + 2\varepsilon_3.$$

Аналогично

$$\Phi'(a) = f'(a+k) - f'(a) = f''(a)\,k + \varepsilon_4,$$

поэтому

$$\Phi'(a+\theta h) - \Phi'(a) = 3\varepsilon_5.$$

и, таким образом,

$$\sup_{0<\theta<1} |\Phi'(a+\theta h) - \Phi'(a)|\,|h| \le 3|\varepsilon_5|\,|h| < 3\varepsilon\delta^{-2}$$

С другой стороны,

$$\Phi'(a)h = f''(a)\,kh + \varepsilon\delta^{-2}$$

поэтому справедливо соотношение

$$|f(a+h+k) - f(a+h) - f(a+k) + f(a) - f''(a)\,kh| \le 4\varepsilon\delta^{-2}$$

Меняя местами h и k, получаем

$$|f''(a)hk - f''(a)kh| \le 8\varepsilon\delta^2.$$

Неравенство (3) имеет место при всех $|h| \le \frac{\delta}{2}$, $|k| \le \frac{\delta}{2}$, (3)
поэтому и

$$\sup_{h,k} |f''(a)hk - f''(a)kh| \le 8\varepsilon\delta^2,$$

Для произвольных $h_0 \in X$, $k_0 \in X$ получаем следовательно

$$\left| f''(a)\frac{\delta}{2}\frac{h_0}{|h_0|}\frac{\delta}{2}\frac{k_0}{|k_0|} - f''(a)\frac{\delta}{2}\frac{k_0}{|k_0|}\frac{\delta}{2}\frac{h_0}{|h_0|} \right| \le 8\varepsilon\delta^2,$$

или, что то же

$$|f''(a)h_0k_0 - f''(a)k_0h_0| \le 32\varepsilon|h_0|\cdot|k_0|.$$

Так как здесь $\varepsilon > 0$ произвольно, то

$$f''(a)h_0k_0 - f''(a)k_0h_0 = 0,$$

что и требовалось.

6. Симметрия смешанных производных. Для дважды дифференцируемой функции $y = f(x)$ $(G \subset X \to Y)$ заданной в области G прямой суммы $X = X_1 + X_2$ пространств X_1 и X_2 определены (2.26) частные производные второго порядка

$$\frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \quad \text{и} \quad \frac{\partial^2 f(x)}{\partial x_2 \partial x_1},$$

совпадающие с результатами соответствующих сужений оператора $f''(x)$ (2.26).

Для этих операторов имеют смысл выражения

$$\frac{\partial^2 f(x)}{\partial x_1 \partial x_2} h_2 h_1 \quad \text{и} \quad \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} h_1 h_2, \quad h_1 \in X_1, h_2 \in X_2,$$

представляющие собой векторы пространства Y.

Поскольку в силу а имеем

$$f''(x)h_2h_1 = f''(x)h_1h_2$$

для любых h_1 и h_2 из X, то, в частности, получаем равенство

$$\frac{\partial^2 f(x)}{\partial x_1 \partial x_2} h_2 h_1 = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} h_1 h_2$$

(4)

для рассматриваемых $h_1 \in X_1$ и $h_2 \in X_2$.

в. Для функции $f(x_1, x_2): R_2 \to Y$ числовых переменных x_1 и x_2 $(X_1 = X_2 = R_1)$ операторы $\frac{\partial^2 f(x)}{\partial x_1 \partial x_2}$ и $\frac{\partial^2 f(x)}{\partial x_2 \partial x_1}$ принимают в каждой точке $x = \{x_1, x_2\}$ значения в том же пространстве Y, и равенство (4) показывает, что эти значения совпадают:

$$\frac{\partial^2 f(x)}{\partial x_1 \partial x_2} = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1}. \tag{5}$$

Равенство (4) имеет место во всяком случае, если существует $f''(x)$, т.е. (3.13) если функции $\frac{\partial f(x)}{\partial x_1}$, $\frac{\partial f(x)}{\partial x_2}$ являются дифференцируемыми. Это условие включает в себя, в частности, существование производных $\frac{\partial^2 f(x)}{\partial x_1^2}$ и $\frac{\partial^2 f(x)}{\partial x_2^2}$.

В классическом анализе равенство (5) устанавливалось при условиях, которые использовали только свойства $\frac{\partial^2 f(x)}{\partial x_1 \partial x_2}$ и $\frac{\partial^2 f(x)}{\partial x_2 \partial x_1}$, без привлечения других вторых производных; затем требовалось существование этих производных в окрестности точки a и непрерывность в самой точке a; такого рода предположений в общей теореме а мы не делаем.

г. С и м м е т р и я в ы с ш и х п р о и з в о д н ы х. Пусть функция $y = f(x)$ $(G \subset X \to Y)$ является p-дифференцируемой $(p \geq 2)$ при $x = a$; покажем, что форма

$$f^{(p)}(a)\, h_p \dots h_1, \quad h_1, \dots, h_p \in X,$$

симметрична. Предположим, что аналогичное утверждение справедливо для всякой $(p-1)$-дифференцируемой функции. Тогда мы можем написать

$$f^{(p)}(a)\, h_p \dots h_1 = g^{(p-1)}(a)\, h_{p-1} \dots h_1, \tag{6}$$

где $g(x) = f'(x) h_p$ есть $(p-1)$-дифференцируемая функция при $x = a$. Поэтому в левой части (6) используя предположение индукции, можно переставлять аргументы $h_{p-1}, \dots, h_1$. Но можно написать и так:

$$f^{(p)}(a)\, h_p \dots h_1 = \omega'(a)\, h_1,$$

где $w(x) = f^{(p-1)}(x) h_p \dots h_2$ есть I-дифференцируемая функция при $x = a$, а в левой части (6) таким образом можно переставлять любые аргументы $h_p \dots h_2$. Так как $p > 2$, то в итоге допустима любая перестановка векторов $h_1, \dots, h_p$, что и требуется.

3.23. **Высшие производные от обобщенного произведения.** Пусть, как в I.346, в области G пространства T заданы дифференцируемые функции $x(t) : G \to X$ и $y(t) : G \to Y$ и их обобщенное произведение $z(t) = \langle x(t), y(t) \rangle : G \to Z$. Теперь мы предположим, что эти функции p раз дифференцируемы, и докажем, что функция $z(t)$ также p раз дифференцируема.

Для первой производной произведения $\langle x(t), y(t) \rangle$ мы получили в I.346 формулу

$$z'(t) = \langle x'(t), y(t) \rangle + \langle x(t), y'(t) \rangle,$$

причем операторы в правой части имеют указанный там смысл:

$$z'(t)\,dt = \langle x'(t)\,dt, y(t) \rangle + \langle x(t), y'(t)\,dt \rangle.$$

Тем самым эти операторы являются снова обобщенными произведениями и в случае $p > 2$, по доказанному, снова дифференцируемы по t. Используя формально метод записи I.346, находим, что

$$z''(t) = \langle x''(t), y(t) \rangle + \langle x'(t), y'(t) \rangle_1 + \langle x'(t), y'(t) \rangle_2 + \langle x(t), y''(t) \rangle. \qquad (1)$$

Индексы 1 и 2 у средних слагаемых поставлены потому, что эти слагаемые имеют различный смысл, в чем нетрудно убедиться, используя запись в дифференциалах:

$$z''(t) h_2 h_1 = \langle x''(t) h_2 h_1, y(t) \rangle + \langle x'(t) h_2, y'(t) h_1 \rangle + \langle x'(t) h_1, y'(t) h_2 \rangle + \langle x(t), y''(t) h_2 h_1 \rangle \qquad (2)$$

Все четыре слагаемых в правой части (1) снова являются обобщенными произведениями, что позволяет при $p > 2$ продолжить дифференцирование. После p шагов мы получим формулу

$$\zeta^{(p)}(x) = \langle x^{(p)}(x), y(x)\rangle + \langle x^{(p-1)}(x), y'(x)\rangle_1 + \dots +$$

$$+ \dots + \langle x^{(p-1)}(x), y'(x)\rangle_p + \langle x^{(p-2)}(x), y''(x)\rangle_1 + \dots +$$

$$+ \dots + \langle x^{(p-2)}(x), y''(x)\rangle_{p(p-1)} + \dots + \langle x(x), y^{(p)}(x)\rangle. \qquad (3)$$

Здесь одинаково записанные слагаемые с различными индексами имеют различный смысл, выявляющийся при записи в дифференциалах: $\zeta^{(p)}(x)\, h_p \dots h_1 = \langle x^{(p)}(x)\, h_p \dots h_1, y(x)\rangle +$

$$+ \langle x^{(p-1)}(x)\, h_p \dots h_2, y'(x)\, h_1\rangle + \dots$$

$$\dots + \langle x^{(p-1)}(x)\, h_{p-1} \dots h_1, y'(x)\, h_p\rangle +$$

$$+ \langle x^{(p-2)}(x)\, h_p \dots h_3, y''(x)\, h_2 h_1\rangle + \dots$$

$$\dots + \langle x^{(p-2)}(x)\, h_{p-2} \dots h_1, y''(x)\, h_p h_{p-1}\rangle + \dots$$

$$\dots + \langle x(x), y^{(p)}(x)\, h_p \dots h_1\rangle. \qquad (4)$$

Если же образовать соответствующую степенную форму, полагая $h_1 = \dots = h_p = h$, то мы получим более простую формулу:

$$\zeta^{(p)}(x)\, h \dots h = \langle x^{(p)}(x) \underbrace{h \dots h}_{p\ раз}, y(x)\rangle +$$

$$+ p\langle x^{(p-1)}(x) \underbrace{h \dots h}_{p-1\ раз}, y'(x)\, h\rangle + \frac{p(p-1)}{2}\langle x^{(p-2)}(x) \underbrace{h \dots h}_{p-2\ раз} y''(x)\, h h\rangle + \dots$$

$$\dots + \langle x(x), y^{(p)}(x) \underbrace{h \dots h}_{p\ раз}\rangle, \qquad (5)$$

или, в символической записи,

$$\zeta^{(p)}(x) = \langle x^{(p)}(x), y(x)\rangle + p\langle x^{(p-1)}(x), y'(x)\rangle +$$

$$+ \frac{p(p-1)}{2}\langle x^{(p-2)}(x), y''(x)\rangle + \dots + \langle x(x), y^{(p)}(x)\rangle \qquad (5')$$

(формула Лейбница); следует, однако, иметь в виду, что смысл эта формула имеет только в реализациях (4) и (5).

3.24. Высшие производные сложной функции. Теорема. Пусть $y=y(x)$ $(G\subset X\to Y)$ - функция, p раз дифференцируемая при $x=a$, и $z=z(y)$ $(W\subset Y\to Z)$ - функция, p раз дифференцируемая при $y=b=f(a)\in W$, тогда сложная функция $z[y(x)]=\zeta(x)$ $(U\subset X\to Z)$ p раз дифференцируема при $x=a$.

Доказательство. При $p=1$ теорема была доказана в I.33. Предполагая теорему справедливой для номера $p-1\geqslant 0$, докажем ее для номера p. Первая производная функции $\zeta(x)$, согласно I.33 имеет вид

$$\zeta'(x)=z'[y(x)]y'(x).$$

Первый множитель есть композиция p раз дифференцируемой функции $y(x)$ и $p-1$ раз дифференцируемой функции $z'(y)$; согласно предположению индукции этот множитель есть $p-1$ раз дифференцируемая функция от x. Второй множитель очевидно, есть $p-1$ раз дифференцируемая функция от x. В силу 3.23 все произведение есть $p-1$ раз дифференцируемая функция от x. Отсюда $\zeta(x)$ есть p раз дифференцируемая функция от x, что и требовалось.

3.25. Высшие производные обратного оператора. Пусть, как в I.35 $x\in L(U,V)$ - обратимый линейный оператор, действующий из пространства U в пространство V, а $x^{-1}:V\to U$ - обратный к x оператор. Покажем, что функция x^{-1} имеет производные по x любого порядка p.

Для $p=1$ это было доказано в I.35б, причем была получена формула

$$d(x^{-1})=-x^{-1}hx^{-1}. \qquad (I)$$

Саму производную $(x^{-1})'$ можно записать в виде обобщенного произведения

$$(x^{-1})'=-\langle x^{-1},x^{-1}\rangle,$$

которое понимать следует, конечно, в смысле (I). Если предположить, что функция x^{-1} дифференцируема p раз, то ее

лесно 3.23, и функция $(x^{-1})'$ будет дифференцируема p раз, т.е. x^{-1} дифференцируема $p+1$ раз. Так как для $p=1$ утверждение справедливо, то приведенное рассуждение показывает, что оно справедливо для любого $p=1,2,\ldots$, что и требовалось.

3.26. В ы с ш и е п р о и з в о д н ы е н е я в н о й ф у н к ц и и.

а. Т е о р е м а. Пусть выполнены условия теоремы о неявной функции 2.13; задана функция $z=\phi(x,y)$ $(V=\{x\in X, y\in Y: |x-a|<\varepsilon, |y-b|<\rho\}\to Z)$, причем $\phi(a,b)=0$ и оператор $\frac{\partial\phi(x,y)}{\partial y}$ непрерывен по x и y и обратим при $x=a$, $y=b$. Тогда, если функция $\phi(x,y)$ p раз дифференцируема в окрестности V, то и неявная функция $y(x)$, являющаяся решением уравнения $\phi(x,y(x))=0$, $y(a)=b$, и существующая в силу теоремы 2.13 также p раз дифференцируема.

Доказательство. Проведем индукцию по номеру q, принимающему значения $1,\ldots,p$, т.е. предполагаем, что функция $y(x)$ дифференцируема $q-1$ раз $(0\le q-1\le p-1)$; докажем, что она дифференцируема q раз. Для $q-1=0$, т.е. $q=1$, утверждение следует из 2.14; там же была получена формула

$$y'(x)=-\left[\frac{\partial\phi(x,y)}{\partial y}\right]^{-1}\cdot\frac{\partial\phi(x,y)}{\partial x}.$$

Первый множитель есть композиция $q-1$ раз дифференцируемой (по предположению индукции и по 3.21б) функции $x\to\{x,y(x)\}$, $p-1$ раз дифференцируемой функции $\frac{\partial\phi(x,y)}{\partial y}$ и бесконечно дифференцируемой функции взятия обратного оператора (3.25); второй множитель есть композиция той же функции $x\to\{x,y(x)\}$ и $p-1$ раз дифференцируемой функции $\frac{\partial\phi(x,y)}{\partial x}$. По 3.24 оба множителя $q-1$ раз дифференцируемы, по 3.23 $q-1$ раз дифференцируемо и их произведение.

Отсюда $y'(x)$ является $q-1$ раз дифференцируемой, а $y(x)$ — q раз дифференцируемой функцией, что и требовалось.

б. В частности, обратная функция $y = \psi(x)$, определяемая из уравнения $x = \varphi(y)$, $a = \varphi(b)$ с обратимым оператором $\varphi'(b)$ (2.17), p раз дифференцируема, если p раз дифференцируема функция $\varphi(y)$.

3.27. а. В 2.43 мы свели теорему о дифференцируемости по параметру λ неподвижной точки сжимающего отображения $A(u, \lambda)$ $(U \times \Lambda \to U)$ к теореме 2.14 о дифференцируемости неявной функции. Используя вместо теоремы 2.42 теорему 3.26 а, получаем следующее усиление теоремы 2.43 на случай высших производных: если правая часть дифференциального уравнения

$$\frac{dy(x,\lambda)}{dx} = \phi(x, y(x,\lambda), \lambda) \qquad (1)$$

и начальное условие

$$y(x_0, \lambda) = \varphi(\lambda) \qquad (2)$$

являются p-дифференцируемыми функциями параметра λ, то и решение $y = y(x, \lambda)$ уравнения (1) с условием (2) является p-дифференцируемой функцией от λ.

б. Аналогичными свойствами дифференцируемости по λ обладают и производные от решения $y(x, \lambda)$. Так, если функция $\phi(x, y, \lambda)$ является в рассматриваемой области p раз дифференцируемой функцией от аргумента (y, λ), то (по 3 а в силу самого уравнения (I)) функция $\frac{dy(x,\lambda)}{dx}$ является p раз дифференцируемой по λ. То же относится к дальнейшим производным от решения $y(x, \lambda)$, уравнения для которых получаются дифференцированием по x уравнения (I) в предположении соответствующей гладкости функции $\phi(x, y, \lambda)$.

§ 3.3. Вариационное исчисление.

3.31а. Пусть T есть отрезок вещественной оси $\alpha \le t \le \beta$, и Y - банахово пространство. Рассмотрим пространство $Y^1(T)$ из всех функций $y = y(t)$, определенных для $t \in T$, принимающих значения в пространстве Y, непрерывных и обладающих непрерывными производными первого порядка, нормированное по формуле

$$\|y\| = \sup_{t \in T} \{ |y(t)|_Y + |y'(t)|_Y \} \qquad (1)$$

Это пространство - полное на основании I.45. Пусть далее $V_c = \{ y \in Y : |y - b| \le c \}$ - шар; совокупность $V_c^1(T)$ всех функций $y \in Y^1(T)$, для которых $y(t) \in V_c$ при каждом $t \in T$, образует в $Y^1(T)$ замкнутое множество. Пусть, наконец, задана C^1 - функция $F(t, y, z) : T \times V_c \times Y \to R_1$.

Ей можно поставить в соответствие функционал $\Phi[y]$, действующий из $V_c^1(T)$ в R_1 по формуле

$$y \in Y^1(T) \longrightarrow \Phi(y) = \int_\alpha^\beta F(t, y(t), y'(t))\, dt. \qquad (2)$$

Нас будут интересовать локальные экстремумы этого функционала, абсолютные и условные.

6. Покажем, прежде всего, что функционал $\Phi(y)$ при указанных условиях является C^1-функционалом в $V_c^1(T)$, т.е. имеет непрерывную производную по y. Для этого достаточно показать, что отображение $F(t, y(t), y'(t)) : V_c^1(T) \to R_1$ имеет непрерывную производную. При каждых фиксированных $t \in T$, $y \in V_c$, $z \in Y$ выражение $F(t, y, z)$ есть число, производная $F_t(t, y, z)$ - число, производные $F_y(t, y, z)$ и $F_z(t, y, z)$ есть операторы из Y в R_1, т.е. линейные непрерывные функционалы в Y - элементы сопряженного пространства Y'. Функции $F_y(t, y, z)$ и $F_z(t, y, z)$ мы будем предполагать равномерно непрерывными в $T \times V_c \times V_\rho$ при любом ρ.

Для заданного $\varepsilon > 0$ найдем $\delta > 0$ так, чтобы из неравенств $|h| \le \delta$, $|k| \le \delta$ при всех $t \in T$, $y \in V_c$, $z \in V_\rho$ следовало

$$\sup_{t \in T} |F_y(t, y + h, z + k) - F_y(t, y, z)| \le \varepsilon;$$

$$\sup_{t \in T} |F_z(t, y + h, z + k) - F_z(t, y, z)| \le \varepsilon;$$

тогда по теореме о конечном приращении

$$|F(x,y+h,z+k)-F(x,y,z)-F_y(x,y,z)h-F_z(x,y,z)k| \le$$

$$\le \sup_{0<\Theta<1} |[F_y(x,y+\Theta h,z+\Theta k)-F_y(x,y,z)]h| +$$

$$+ \sup_{0<\Theta<1} |[F_z(x,y+\Theta h,z+\Theta k)-F_z(x,y,z)]k| \le \varepsilon(|h|+|k|). \quad (3)$$

Для выяснения вопроса о дифференцируемости функционала $\Phi[y]$ мы должны рассмотреть приращение этого функционала, т.е. величину $\Phi[y+h]-\Phi[y]$ (где $h=h(x)\in Y(T)$), и выделить из этого приращения главную линейную часть.

Мы имеем

$$\Phi[y+h]-\Phi[y]=\int_\alpha^\beta [F(x,y(x)+h(x),y'(x)+h'(x))-F(x,y(x),y'(x))]\,dx.$$

Подставляя в (3) вместо y и h значения $y(x)$ и $h(x)$, вместо z и k значения $y'(x)$ и $h'(x)$, мы получаем

$$F(x,y(x)+h(x),y'(x)+h'(x))-F(x,y(x),y'(x))=$$

$$=F_y(x,y(x),y'(x))h(x)+F_z(x,y(x),y'(x))h'(x)+\varepsilon(x),$$

где $|\varepsilon(x)| \le \varepsilon(|h(x)|+|h'(x)|) \le \varepsilon\|h\|$.

Отсюда $\Phi[y+h]-\Phi[y]=$

$$=\int_\alpha^\beta [F_y(x,y(x),y'(x))h(x)+F_z(x,y(x),y'(x))h'(x)]\,dx+R,$$

где $|R|=\left|\int_\alpha^\beta \varepsilon(x)\,dx\right| \le \varepsilon(\beta-\alpha)\|h\|$

Первый член в ( ) представляет собой линейное выражение относительно $h(x)\in Y^1(T)$, второй – малую высшего порядка. Отсюда следует, что функционал $\Phi(y)$ есть дифференцируемая функция из $Y^1(T)$ в R_1 с дифференциалом (иначе называемым вариацией; отсюда – вариационное исчисление)

$$\Phi'(y)h=\int_\alpha^\beta [F_y(x,y(x),y'(x))h(x)+F_z(x,y(x),y'(x))h'(x)]\,dx. \quad (4)$$

(3)

3.32. Будем разыскивать точки локального экстремума функционала $\Phi(y)$. В этих точках величина $\Phi'(y)h$ должна обращаться в нуль для любого смещения $h \in Y'(T)$. Будем сначала брать смещения $h(x)$ такие, что $h(\beta) = h(\alpha) = 0$. Выражение $\Phi'(y)h$ можно преобразовать интегрированием по частям, либо во втором слагаемом переходя от $h'(x)$ к $h(x)$, либо в первом слагаемом - от $h(x)$ к $h'(x)$. Первый прием, однако, требует дифференцируемости по x функции $F_z(x, y(x), y'(x))$, которая, вообще говоря, при произвольной $y(x) \in V'(T)$ не имеет места. Поэтому применим второй прием. Будем обозначать первообразную от функции $\varphi(x) \in V'(T)$ через $I\varphi(x)$, тогда мы получим

$$\int_\alpha^\beta F_y(x, y(x), y'(x))\,dx = I F_y(x, y(x), y'(x))\,h(x)\Big|_\alpha^\beta -$$

$$- \int_\alpha^\beta I F_y(x, y(x), y'(x))\,h'(x)\,dx = - \int_\alpha^\beta I F_y(x, y(x), y'(x))\,h'(x)\,dx;$$

откуда

$$\Phi'(y)h = \int_\alpha^\beta \left(- I F_y(x, y(x), y'(x)) + F_z(x, y(x), y'(x)\right) h'(x)\,dx$$

Это выражение должно обращаться в нуль при всякой функции $h(x) \in Y'(T)$ такой, что $h(\alpha) = h(\beta) = 0$.

6. Лемма. (лемма Дю-Буа-Раймонда). Пусть $T = \{\alpha \le x \le \beta\}$, Y -банахово пространство и Y' - сопряженное к Y. Если для некоторой непрерывной функции $\varphi(x) : T \to Y'$ и для любой функции $h(x) \in Y(T)$, $h(\alpha) = h(\beta) = 0$, выполняется равенство

$$\int_\alpha^\beta \varphi(x)\, h'(x)\,dx = 0,$$

то $\varphi(x) = const$.

Доказательство. Положим

$$c = \frac{1}{\beta - \alpha} \int_\alpha^\beta \varphi(x)\,dx \in Y'$$

и покажем, что $\varphi(x) \equiv c$. Для этого, в силу лем-

2.31 б достаточно показать, что для любой непрерывной функции

$$k(t) \in Y(T) \qquad \int_{\alpha}^{\beta} [\varphi(t) - c]\, k(t)\, dt = 0.$$

Для данной функции $k(t)$ положим

$$\alpha = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} k(t)\, dt \in Y$$

и $p(t) = k(t) - \alpha$; очевидно, что

$$\int_{\alpha}^{\beta} p(t)\, dt = \int_{\alpha}^{\beta} k(t)\, dt - \alpha(\beta - \alpha) = 0.$$

Поэтому

$$\int_{\alpha}^{\beta} [\varphi(t) - c]\, k(t)\, dt = \int_{\alpha}^{\beta} [\varphi(t) - c]\, p(t)\, dt + \int_{\alpha}^{\beta} [\varphi(t) - c]\, \alpha\, dt$$

Но если положить $h(t) = \int_{\alpha}^{t} p(\xi)\, d\xi$, мы будем иметь

$$h(t) \in Y'(T),\ h'(t) = p(t),\ h(\alpha) = 0,\ h(\beta) = \int_{\alpha}^{\beta} p(\xi)\, d\xi = 0;$$

поэтому, в силу условия леммы

$$\int_{\alpha}^{\beta} [\varphi(t) - c]\, p(t)\, dt = \int_{\alpha}^{\beta} [\varphi(t) - c]\, h'(t)\, dt =$$

$$= \int_{\alpha}^{\beta} \varphi(t)\, h'(t)\, dt - c \int_{\alpha}^{\beta} h'(t)\, dt = 0;$$

также и $\int_{\alpha}^{\beta} (\varphi(t) - c)\, \alpha\, dt = \int_{\alpha}^{\beta} [\varphi(t) - c]\, dt \cdot \alpha = 0.$

Поэтому для любой $k(t) \in Y(T)$ левая часть в (I) равна 0, чего, как мы видели, достаточно для доказательства леммы.

в. Применяя лемму к рассматриваемому равенству 3.3I (4), мы делаем вывод: в искомой стационарной точке выполняется соотношение

$$I\ F_y(t, y(t), y'(t)) - F_z(t, y(t), y'(t)) = const. \qquad (2)$$

Так как первое слагаемое, как первообразная от непрерывной функции, имеет непрерывную производную, то и второе слагаемое $F_z(t, y(t), y'(t))$ обладает этим же свойством. Поэтому равенство (2) можно продифференцировать по t ; мы получим

уравнение

$$\mathcal{F}_y(t, y(t), y'(t)) - \frac{d}{dt}\mathcal{F}_z(t, y(t), y'(t)) = 0 \qquad (3)$$

в котором слагаемые – также непрерывные функции. Уравнение (3) называется уравнением Эйлера.

г. Чтобы глубже проникнуть в структуру уравнения Эйлера наложим на функцию $\mathcal{F}(t, y, z)$ дальнейшие условия. Именно, будем сейчас предполагать, что $\mathcal{F}(t, y, z)$ имеет непрерывные производные по t, y, z не только первого, но и второго порядка. Заметим при этом, что при каждых фиксированных t, y, z, выражение $\mathcal{F}_{tt}(t, y, z)$ есть число, выражения $\mathcal{F}_{ty}, \mathcal{F}_{tz}$ есть линейные функционалы над Y, а выражения $\mathcal{F}_{yy}, \mathcal{F}_{yz}, \mathcal{F}_{zz}$ есть линейные операторы из Y в Y'; для этих последних определены числовые билинейные формы $\mathcal{F}_{yy}hk$, $\mathcal{F}_{yz}hk$, $\mathcal{F}_{zz}kk$, где h и k любые элементы из Y. Будем предполагать далее, что оператор $\mathcal{F}_{zz}: Y \to Y'$ обратим, существует (и непрерывен) оператор $\mathcal{F}_{zz}^{-1}: Y' \to Y$. Применяя к $\mathcal{F}_z(t, y, z)$ формулу конечных приращений, находим

$$\left| \mathcal{F}_z(t+\tau, y+h, z+k) - \mathcal{F}_z(t, y, z) - \mathcal{F}_{zt}(t, y, z)\tau - \mathcal{F}_{zy}(t, y, z)h - \right.$$

$$\left. - \mathcal{F}_{zz}(t, y, z)k \right| \le \sup_{\theta} \left| \mathcal{F}_{zt}(t+\theta\tau, y+\theta h, z+\theta k) - \mathcal{F}_{zt}(t, y, z) \right| \cdot |\tau| +$$

$$+ \sup_{\theta} \left| \mathcal{F}_{zy}(t+\theta\tau, y+\theta h, z+\theta k) - \mathcal{F}_{zy}(t, y, z) \right| \cdot |h| +$$

$$+ \sup_{\theta} \left| \mathcal{F}_{zz}(t+\theta\tau, y+\theta h, z+\theta k) - \mathcal{F}_{zz}(t, y, z) \right| \cdot |k| =$$

$$= o(|\tau|) + o(|h|) + o(|k|)$$

В силу предположенной непрерывности $\mathcal{F}_{zt}$, $\mathcal{F}_{zy}$, $\mathcal{F}_{zz}$. Подставим сюда вместо y значение $y(t)$, вместо z – значение $y'(t)$, вместо h – значение $y(t+\tau) - y(t)$ и вместо k – значение $y'(t+\tau) - y'(t)$. Мы получим

$$|F_z(x+\tau, y(x+\tau), y'(x+\tau)) - F_z(x, y(x), y'(x)) - F_{zx}(x, y(x), y'(x))\tau -$$

$$- F_{zy}(x, y(x), y'(x))(y(x+\tau)-y(x)) - F_{zz}(x, y(x), y'(x))(y'(x+\tau)-y'(x))| =$$

$$= o(|\tau|) + o(|y(x+\tau)-y(x)|) + o(|y'(x+\tau)-y'(x)|) \qquad (4)$$

Из этого соотношения мы выведем сейчас дифференцируемость функции $y'(x)$. В правой части первые два слагаемых имеют вид $o(\tau)$, ввиду предположенной дифференцируемости $y(x)$. Последнее слагаемое, имеет вид $o(\Delta y')$, где

$$\Delta y' = y'(x+\tau) - y'(x).$$

В левой части первая разность допускает выделение главной линейной части по τ в силу доказанной дифференцируемости функции $F_z(x, y(x), y'(x))$ по x. То же относится к следующим двум слагаемым. Таким образом для последнего слагаемого получается равенство вида

$$F_{zz}(x, y(x), y'(x))[y'(x+\tau) - y'(x)] = A\tau + o(\tau) + o(\Delta y').$$

Умножив его на F_{zz}^{-1}, мы получим равенство вида

$$\Delta y' = B\tau + o(\tau) + o(\Delta y') \qquad (5)$$

д. Лемма. Пусть для функции $\varphi(x)$, непрерывной при $x = 0$ выполняется соотношение

$$\Delta\varphi = Bx + o(x) + o(\Delta\varphi). \qquad (6)$$

Тогда имеет место и равенство $\Delta\varphi = Bx + o(x)$,

т.е. функция $\varphi(x)$ при $x = 0$ дифференцируема и $\varphi'(0) = B$.

Доказательство. Мы имеем из (6)

$$|\Delta\varphi| \leq \|B\| \cdot |x| + |o(x)| + |o(\Delta\varphi)|$$

При достаточно малом x имеем $|o(\Delta\varphi)| < \frac{1}{2}|\Delta\varphi|$, $|o(x)| \leq |x|$, отсюда при этих x

$$\frac{1}{2}|\Delta\varphi| \leq \|B\| \cdot |x| + |o(x)| \leq (\|B\|+1) \cdot |x|$$

так что $|\Delta\varphi| \leq C|x|$; для заданного $\varepsilon > 0$ найдем $\delta > 0$ так, чтобы при $|x| < \delta$ иметь $|o(x)| < \varepsilon|x|$, $|o(\Delta\varphi)| < \varepsilon|\Delta\varphi|$; тогда мы получим $\Delta\varphi = Bx + R$,

где $|R| \leq |o(\alpha)| + |o(\Delta\varphi)| \leq \varepsilon|\alpha| + \varepsilon|\Delta\varphi| \leq \varepsilon(1+c)|\alpha|$ имеет порядок малости высшей по сравнению с α . Это означает, что $\varphi(\alpha)$ дифференцируема при $\alpha = 0$ и $\varphi'(0) = B$, что и требуется.

е. Применяя доказанную лемму к функции $y'(t)$ и используя равенство (4), мы приходим к дифференцируемости функции $y'(t)$. Иначе говоря, искомая стационарная функция $y(t)$ имеет не только первую, но и вторую производную. Поэтому уравнение (3) преобразуется к виду

$$F_y(t, y(t), y'(t)) - F_{zt}(t, y(t), y'(t)) - F_{zy}(t, y(t), y'(t)) y'(t) -$$

$$- F_{zz}(t, y(t), y'(t)) y''(t) = 0 \qquad (7)$$

Таким образом, уравнение Эйлера приведено к виду обыкновенного дифференциального уравнения 2-го порядка. Его решения называются экстремалями функционала $\Phi(y)$. В силу основной теоремы существования и единственности через каждую точку $y_0 \in V_\varepsilon$ в каждом направлении $y_0' \in Y$ приходит единственная экстремаль – в дальнейшем предположим, что функции F_y, F_{zt}, F_{zy} и F_{zz}^{-1} хотя бы удовлетворяют условию Липшица по аргументам y и z .

3.33а. Многие задачи вариационного исчисления формулируются так: среди всех кривых $y = y(t) \in Y^1(T)$ с заданными значениями $y(\alpha)$ и $y(\beta)$ найти такую, для которой функционал 3.31 (2) принимает экстремальное значение. Для решения этой задачи, как мы теперь видим, нужно выделить экстремаль, т.е. решение уравнения Эйлера 3.32 (3) или 3.32 (7), проходящую через точки $(\alpha, y(\alpha))$ и $(\beta, y(\beta))$. Встречаются и задачи "со свободными концами", где значения $y(\alpha)$ и $y(\beta)$ не предписаны, о них мы поговорим дальше. Заметим сейчас, что и задача "с закрепленными концами", т.е. с заданными $y(\alpha)$ и $y(\beta)$ есть нелокальная задача для уравнения 3.32 (7), и она не всегда имеет решение даже при выполнении всех условий, обеспечивающих локальную разрешимость уравнения 3.32 (7). Рассмотрим сейчас

несколько частных случаев и примеры.

б. $\mathcal{F}$ зависит только от z.

В этом случае уравнение Эйлера 3.32 (7) приобретает вид

$$\mathcal{F}_{zz}(z)\cdot y''(t)=0$$

откуда, в том же предположении обратимости $\mathcal{F}_{zz}$, и

$$y''(t)=0;$$

это означает, что $y(t)$ - линейная функция:

$$y(t)=k(t-\alpha)+y_1,\ y_1,k\in Y.$$

Пусть $Y=R_n$ и $Y\times T=R_{n+1}$ с обычной евклидовой метрикой; известно, что длина дуги кривой $y=y(t)$ между точками $\{\alpha,y_1\}$ и $\{\beta,y_2\}$ выражается формулой

$$S=\int_{\alpha}^{\beta}\sqrt{1+|y'(t)|^2}\,dt$$

Найдем кривую минимальной длины, соединяющей эти точки. Так как здесь $F(z)=\sqrt{1+|z|^2}$ зависит только от z, то можно применить полученный результат: решение изображается линейной функцией, т.е. является прямой линией.

в. F зависит от t и z, но не зависит от y: $\mathcal{F}=\mathcal{F}(t,z)$. В этом случае уравнение Эйлера 3.32 (3) может быть записано в виде $\frac{d}{dt}\mathcal{F}_z(t,y'(t))=0$

откуда

$$\mathcal{F}_z(t,y'(t))=const;$$

для функции $y(t)$ получается уравнение I-го порядка, содержащее произвольную постоянную.

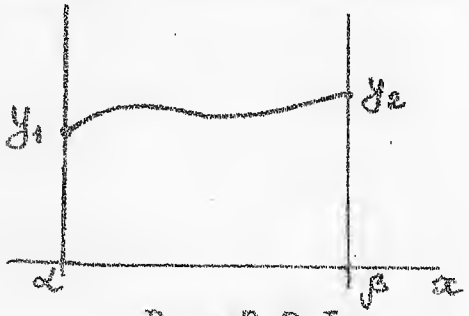

Рис. 3.3-I

Найдем в $T\times R_1$ с обычной евклидовой метрикой минимальный по времени путь из точки $\{a,y_1\}$ в точку $\{b,y_2\}$ при условии, что в точке с абсциссой x скорость движения равна заданной функции $v(x)$. Мы имеем здесь для элемента dt времени t выражение

$$dt = \frac{dS}{v} = \frac{\sqrt{dx^2+dy^2}}{v(x)} = \frac{\sqrt{1+y'^2(x)}\,dx}{v(x)},$$

а для всего времени T - результат

$$T = \int_a^b \frac{\sqrt{1+y'^2(x)}\,dx}{v(x)}$$

В соответствии со сказанным уравнение Эйлера преобразуется к виду

$$F_2(x,y,y') = \frac{y'}{\sqrt{1+y'^2}\cdot v(x)} = c, \qquad (I)$$

Величина $\frac{y'}{\sqrt{1+y'^2}}$ есть синус ведущего угла, т.е. угла между касательной к кривой $y = y(x)$ и осью x. Мы видим, что искомый оптимальный путь должен подчиняться закону: синус ведущего угла должен быть пропорционален скорости в данной точке пути. Например, если $v(x)$ имеет одно постоянное значение v_1 при $a \le x \le c$ и другое постоянное значение v_2 при $c \le x \le b$, то на каждом из участков (a,c) и (c,b) и ведущий угол постоянен, т.е. движение совершается по прямой; в точке пересечения этих отрезков выполняется соотношение

$$\frac{v_2}{v_1} = \frac{\sin\varphi_2}{\sin\varphi_1}$$

что представляет собой закон преломления света в форме Гюйгенса-Френеля.

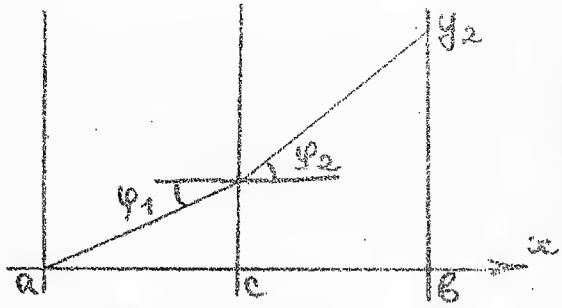

Рис. 3.3-2

В общем случае, разрешая уравнение (I) относительно y', получаем

$$y' = \frac{c\,v(x)}{\sqrt{1-c^2v^2(x)}}$$

Уравнения экстремалей имеют вид

$$y = y(x) = \int_a^x \frac{c\,v(\xi)\,d\xi}{\sqrt{1-c^2v^2(\xi)}} + y_1$$

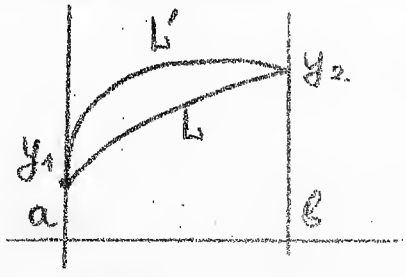

Рис. 3.3-3

Так как $y(b)$ при $c=0$ обращается в y_1 и при c, увеличивающемся до $c_0 = \min(1/v(x))$ увеличивается, то правые концы экстремалей покрывают отрезок вертикальной прямой $x=b$ от значения $y=y_1$ до значения

$$y_2 = \int_a^b \frac{c_0 v(\xi)\,d\xi}{\sqrt{1-c_0^2 v^2(\xi)}} + y_1 \qquad (2)$$

Если этот интеграл бесконечен, то существует экстремаль, ведущая из точки $\{a, y_1\}$ в точку $\{b, y_2\}$ с любым $y_2 > 0$. В частности, так будет при линейном законе $v(\xi)$. Если же интеграл (2) конечен (например, для $v(\xi) = \sqrt{1-(\xi-a)\gamma}$, $\gamma < 1/(b-a)$), экстремали, ведущей в точку $\{b, y_2\}$, при достаточно больших y_2 не существует; это означает, что для любого пути L, ведущего из $\{a, y_1\}$ в $\{b, y_2\}$ есть и более "экономный" по времени путь L, теснее примыкающий к оси y и более эффективно использующий большие скорости при малых $x-a$. (Рис. 3.3-3).

г. F зависит от y и z, но не зависит от x.

Уравнение Эйлера 3.32 (7) принимает вид

$$F_y - F_{zy}\,y' - F_{zz}\,y'' = 0$$

Все слагаемые здесь-линейные функционалы в пространстве y. Применяя их к y' и добавляя и вычитая $F_z y''$, мы получим

$$F_y\,y' + F_z\,y'' - F_{zy}\,y'\,y' - F_{zz}\,y'\,y'' - F_z\,y'' = 0$$

Теперь в правой части стоит производная по x от выражения

$$F(y, y') - F_z(y, y')\,y'$$

Следовательно, на каждой экстремали эта величина есть постоянная:

$$F(y, y') - F_z(y, y')\,y' = c \in Y \qquad (3)$$

Мы привели уравнение Эйлера 3.32 (7) к уравнению I-го порядка (с произвольным элементом $c \in Y$).

Пример. Найдем кривую $y = y(x) : [a, b] \to R_1$ проходящую через точки $\{a, y(a)\} \in R_2$, $\{b, y(b)\} \in R_2$ для которой площадь поверхности вращения вокруг оси x минимальна. Здесь речь идет о минимуме функционала

$$\Phi[y] = 2\pi \int_a^b y\sqrt{1+y'^2}\,dx.$$

Соотношение (3) дает нам

$$y\sqrt{1+y'^2}-y'\frac{yy'}{\sqrt{1+y'^2}}=C\in R_1.$$

Отсюда

$$y'=\sqrt{\left(\frac{y}{C}\right)^2-1}$$

$$\frac{dy}{\sqrt{\left(\frac{y}{C}\right)^2-1}}=dx$$

Полагая $y=C\cdot ch\,t$, $dy=C\cdot sh\,t\,dt$, находим $dx=C\,dt$, откуда $x=C\cdot t+C_1$, так что семейство экстремалей имеет уравнение

$$y=C\cdot ch\,\frac{x-C_1}{C}$$

Все экстремали получаются из кривой $y=ch\,x$ преобразованиями подобия и сдвигами вдоль оси x (см. рис. 3.3-4).

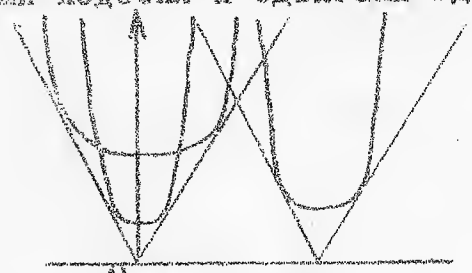

Рис. 3.3-4

Будем считать, что $\alpha=-C$, $\beta=C$, $y(\alpha)=y(\beta)$. Тогда искомая кривая очевидно должна принадлежать к семейству $y=C\cdot ch\,\frac{x}{C}$

В зависимости от соотношения между C и $y(C)$ могут возникнуть три разные возможности:

через заданные точки проходят две экстремали, одна или ни одной (рис. 3.3-5). Если экстремаль, проходящая через точки $(-\alpha,\beta)$, (α,β) имеется, то площадь соответствующей поверхности вращения легко вычислить:

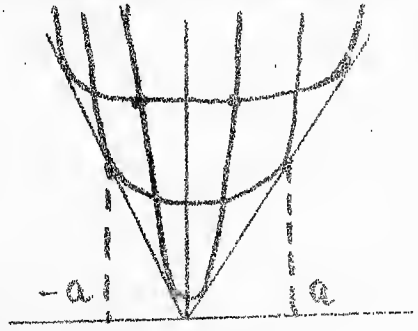

Рис. 3.3-5

откуда ясно, что в случае, когда через две точки проходят две экстремали, меньшее значение площади поверхности вращения осуществляется для той, которая расположена выше.

д. Еще один важный пример связан с механикой системы.

Пусть механическая система описывается координатами $x_1, \dots, x_n$ и скоростями $v_1, \dots, v_n$, которые подчинены уравнениям

$$\left.\begin{aligned} \dot{x}_i &= v_i \\ m_i \dot{v}_i &= -\frac{\partial U}{\partial x_i} \end{aligned}\right\} \quad i = 1, \dots, n$$

Пусть $T = \frac{1}{2}\sum m_i v_i^2$ кинетическая энергия системы и $L = T - U$ ее функция Лагранжа. Фиксируя моменты времени α и β, рассмотрим функционал, называемый действием:

$$S[x(t)] = \int_\alpha^\beta L(t, x(t), x'(t))\,dt$$

где $x(t) = \{x_1(t), \dots, x_n(t)\}$ некоторая C^1-функция ($[\alpha, \beta] \to R_n$). Найдем условие стационарности функционала $\Phi[x(t)]$. На соответствующей линии $x = x(t)$ выполняется уравнение Эйлера

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial v} = 0 \qquad (4)$$

Поскольку $L(t, x, v) = T - U = \frac{1}{2}\sum m_i v_i^2 - U(x)$;

мы имеем $\frac{\partial L}{\partial x} = \left\{-\frac{\partial U}{\partial x_i}\right\}$, $\frac{\partial L}{\partial \dot{x}} = \{m_i \dot{x}_i\}$, и уравнение (4) превращается в систему уравнений Ньютона

$$-\frac{\partial U}{\partial x_i} - m_i \ddot{x}_i = 0$$

Таким образом, движение системы происходит по экстремали функционала действия. Во многих случаях (но не всегда) на этой экстремали реализуется минимум действия (см. ниже, 3.37 б); поэтому полученный результат называют также принципом наименьшего действия.

3.34. Пучки экстремалей.

В рассмотренных примерах видно, что экстремали, вообще говоря, рассматриваемые в большой области проходят не через любую пару точек. Оказывается, что для небольших областей – и при выполнении всех условий 3.32, обеспечивающих существование и единственность траекторий с заданным началом и направлением – задача о проведении экстремали через две заданные точки разрешима. Рассмотрим этот вопрос подробнее.

Совокупность всех экстремалей, $y = y(x)$ функционала $\Phi(y)$, проходящих при $x = \alpha$ через заданную точку $y_0 \in Y$, называется пучком экстремалей. Каждая из них определяется вектором y_0'. Выберем одну из этих экстремалей $y = y(x)$ и положим $y_1 = y(\beta)$. Покажем, сначала, что при достаточно малом $\beta - \alpha$ и достаточно малом $\varepsilon > 0$ экстремали проходят через любую точку (β, y), где $|y - y_1| < \varepsilon$.

Для доказательства запишем уравнение Эйлера 3.32 (7) в форме системы двух уравнений первого порядка

$$\begin{cases} y'(x) = v(x) \\ v'(x) = f(x, y(x), v(x)) \end{cases} \quad \text{или} \quad \begin{cases} y(x) = y(\alpha) + \int_\alpha^x v(\xi)\,d\xi \\ v(x) = v(\alpha) + \int_\alpha^x f(\xi, y(\xi), v(\xi))\,d\xi \end{cases} \tag{I}$$

функция $f(x, y, v)$ предполагается непрерывной и ограниченной в области $0 \leqslant x \leqslant x_1$, $|y - y_0| \leqslant \nu$, $|v - v_0| \leqslant \rho$, вместе с ее частными производными по y и по v.

Положим

$$M = \sup_{\substack{0 \leqslant x \leqslant x_1 \\ |y - y_0| \leqslant \nu \\ |v - v_0| \leqslant \rho}} \left(\left| \frac{\partial f(x, y, v)}{\partial y} \right| + \left| \frac{\partial f(x, y, v)}{\partial v} \right| \right)$$

Нас интересуют решения системы (I) при начальных данных $y(0) = y_0$ (фиксированном) и $v(0) = u$ (меняющемся). Обозначим решение через $\{y(x, u), v(x, u)\}$. Как было показано в 2.43 эти функции обладают непрерывными производными по u в окрестности точки $u = v_0$, например, при $|u - v_0| < \varepsilon$

Положим $$N = \sup_{\substack{0 \leqslant x \leqslant x_1 \\ |u - v_0| \leqslant \varepsilon}} \left(\left| \frac{\partial y(x, u)}{\partial u} \right| + \left| \frac{\partial v(x, u)}{\partial u} \right| \right)$$

Мы желаем показать, что оператор $\frac{\partial y(t,u)}{\partial u}$ $(y \to y)$ при малых t обратим. Мы имеем

$$\frac{\partial v(t,u)}{\partial u} = \frac{\partial}{\partial u}\left(u + \int_0^t f(\xi, y(\xi), v(\xi))\,d\xi\right) =$$

$$= E + \int_0^t \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial u}\right)d\xi = E + B(t),$$

где $$\|B(t)\| = \left\|\int_0^t \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial u}\right)d\xi\right\| \le$$

$$\le \int_0^t \left(\left|\frac{\partial f}{\partial y}\right|\left|\frac{\partial y}{\partial u}\right| + \left|\frac{\partial f}{\partial v}\right|\left|\frac{\partial v}{\partial u}\right|\right)d\xi \le M \cdot N\, t$$

Можно взять t настолько малым, положим, $0 \le t \le t_1 = \frac{1}{2MN}$, что при этих t

$$\|B(t)\| \le \frac{1}{2}$$

Далее, при тех же t, $0 \le t \le t_1$, имеем

$$\frac{\partial y(t,u)}{\partial u} = \frac{\partial}{\partial u}\left(y_0 + \int_0^t v(\xi,u)\,d\xi\right) = \int_0^t \frac{\partial v(\xi,u)}{\partial u}\,d\xi =$$

$$= \int_0^t \left(E + B(\xi)\right)d\xi = t\left(E + \frac{1}{t}\int_0^t B(\xi)\,d\xi\right) \qquad (2)$$

Здесь $$\left\|\frac{1}{t}\int_0^t B(\xi)\,d\xi\right\| \le \frac{1}{t}\int_0^t \|B(\xi)\|\,d\xi \le \frac{1}{2},$$

поэтому оператор в скобках в (2) обратим; а тогда обратим и оператор $\frac{\partial y(t,u)}{\partial u}$ при всяком $t > 0$, $t \le t_1$.

По теореме о неявной функции уравнение $y = y(t,u)$ можно разрешить относительно u в окрестности значения $u = u_0$, например, при $|u - u_0| \le \varepsilon_1 \le \varepsilon$, иначе говоря, по любому значению $\bar{y}$ в некоторой окрестности значения $y_1 = y(t_1, u_0)$, например, при $|\bar{y} - y_1| \le \delta$ и t_1 можно однозначно вос-

становить $\bar{u}$, $|\bar{u}-v_0|\le\varepsilon_1$; так, что траектории $y(t,\bar{u})$ при $t=t_1$ проходит как раз через значение $\bar{y}$. Но это и есть то утверждение, которое нам нужно доказать: пучок экстремалей, проходящих через точку $(0,y_0)$ покрывает при $t=t_1$ целую окрестность точки y_1.

б. Заметим, что размеры t_1 и δ зависят лишь от свойств функции $f(t,y,v)$ в окрестности значений $t=0$, $y=y_0$, $v=v_0$. Если мы начнем построение не в точке $\{0,y_0\}$, а в точке $\{t_1,\bar{y}\}$ в пределах этой же окрестности, и будем проводить построение влево по оси t, мы получим пучок траекторий с теми же свойствами и характеристиками: эти траектории, определенные для коэффициентов u, $|u-v(t_1,v_0)|\le\varepsilon$, покроют для значения $t=0$ в пространстве y шар $|y-y_0|\le\delta$.

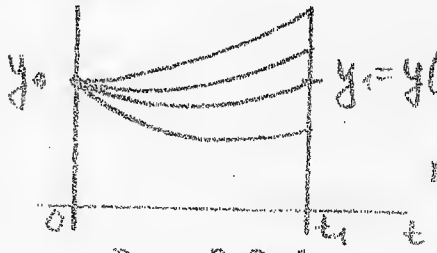

Рис. 3.3-6

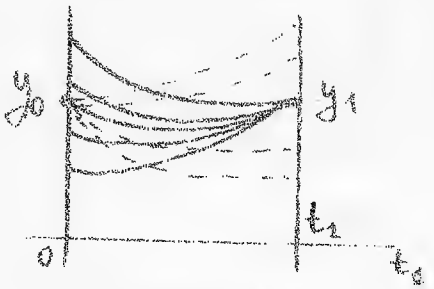

Рис. 3.3-7

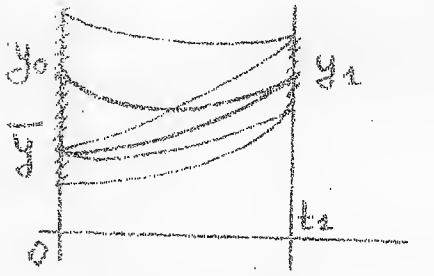

Рис. 3.3-8

в. Будем теперь вместо y_0 в предыдущем построении брать величину $\bar{y}$ из этого шара и в качестве коэффициента v_0 брать коэффициент $\bar{v}$ той траектории, которая ведет из $\bar{y}$ в y_1. Тогда по тем же соображениям траектории $y(t,\bar{v})$ для $|\bar{v}-v|\le\varepsilon_1$ при $t=t_1$ будут покрывать тот же шар $|\bar{y}_1-y_1|\le\delta$. Мы видим, что для любых двух точек $\bar{y}$, $|\bar{y}-y_0|\le\delta$ и $\bar{y}_1$, $|\bar{y}_1-y_1|\le\delta$ существует траектория, и при том единственная, которая ведет из точки $\{0,\bar{y}\}$ в точку $\{t_1,\bar{y}_1\}$.

3.35. Таким образом, хотя и в малых областях, у нас имеются экстремали, соединяющие две заданные точки. На такой экстремали функционал $\Phi(y)$ имеет стационарное значение по сравнению с соседними кривыми, имеющими те же концевые точки. В каких случаях это стационарное значение будет экстремальным-минимальным или

максимальным ? Редко можно ответить на этот вопрос прямым рассмотрением значений функционала на соседних кривых. В общем анализе (3.3) некоторые дальнейшие выводы мы могли делать, анализируя второй дифференциал функции. Найдем второй дифференциал функционала $\Phi[y]$ на экстремали $y = y(x)$. Будем считать, что функция F имеет непрерывные производные до 3-го порядка. Мы имеем в стационарной точке функции $F(x, y, z)$

$$F(x, y+h, z+k) - F(x, y, z) = \frac{1}{2}(F_{yy}h^2 + F_{yz}hk + F_{zz}k^2) + o(|h^2| + |k^3|)$$

Отсюда, как и при вычислении первого дифференциала, мы имеем

$$d^2\Phi[y] = \frac{1}{2}\int_\alpha^\beta [F_{yy}(x, y(x), y'(x))h^2(x) +$$

$$+ F_{yz}(x, y(x), y'(x))h(x)h'(x) + F_{zz}(x, y(x), y'(x))(h'(x))^2]\,dx$$

Второе слагаемое можно преобразовать интегрированием по частям (поскольку существование $y''(x)$ уже известно)

$$\int_\alpha^\beta F_{yz}h(x)h'(x)\,dx = \frac{1}{2}F_{yz}h^2(x)\Big|_\alpha^\beta - \frac{1}{2}\int_\alpha^\beta \frac{d}{dx}F_{yz}h^2(x)\,dx$$

причем внеинтегральный член обращается в нуль ($h(\alpha) = h(\beta) = 0$). Отсюда

$$d^2\Phi[y] = \frac{1}{2}\int_\alpha^\beta [P(x, y(x), y'(x))h^2(x) + F_{zz}(x, y(x), y'(x))[h'(x)]^2]\,dx,$$

где

$$P = F_{yy} - \frac{1}{2}\frac{d}{dx}F_{yz}.$$

Мы утверждаем, что функционал $\Phi[y]$ может иметь на функции $y(x)$ локальный минимум лишь в случае, если при каждом $x \in T$ форма $F_{zz}(x, y(x), y'(x))k^2 \geq 0$

для любого $k \in Y$.

Допустим, что это условие не выполнено, т.е. при некотором $x = x_0$ существует элемент $k \in Y$ такой, что

$F_{zz}(x_0, y(x_0), y'(x_0))k^2 < 0$. Можно считать, что $x_0 \in (\alpha, \beta)$. В силу непрерывности функций F_{zz}, $y(x)$ и $y'(x)$ существует число $\varphi < 0$ и окрестность $(x_0 - \delta, x_0 + \delta) \subset T$, в которой выполняется неравенство $F_{zz}(x, y(x), y'(x))k^2 \leq \varphi < 0$.

Рассмотрим вещественную гладкую функцию $\gamma(t)$, определенную при всех $t \in (-\infty, \infty)$, равную 0 при $|t| \geq \delta'$ и положительную при $|t| < \delta'$, не превосходящую 1, с производной $\gamma'(t)$, большей 1 в интервале $|t| < \delta'' < \delta'$. Функция $\gamma_m(t) = \gamma[m(t-t_0)]$ при любом $m = 1, 2, \ldots$ принадлежит $y^1(T)$. Рассмотрим выражение $d^2\phi[y]$ для смещения $h(t) = \gamma_m(t)k$. Первое слагаемое

$$S_1 = \int_\alpha^\beta P(t, y(t), y'(t))\,\gamma_m^2(t)k^2\,dt$$

остается ограниченным при $m \to \infty$, так как $\gamma_m^2(t) \leq 1$. Второе слагаемое

$$S_2 = \int_\alpha^\beta F_{z'z'}(t, y(t), y'(t))\,[\gamma_m'(t)]^2 k^2\,dt$$

отрицательно и так как на интервале $|t| < \delta''/m$ величина

$$\gamma_m'(t) \geq m, \quad S_2 \leq -q m^2 \cdot \frac{\delta''}{m} = -q m \delta'' \to -\infty$$

Для достаточно больших m величина $d^2\phi(y)$ на смещениям $\gamma_m(t)$ отрицательна. Поэтому, на функции $y(t)$ функционал $\phi[y]$ не может иметь локального минимума. Следовательно, условие $F_{z'z'}(t, y(t), y'(t))\,k^2 \geq 0$ при любом $k \in Y$ и любом $t \in T$ является необходимым для того, чтобы функционал $\phi[y]$ на функции $y(t)$ достигал локального минимума.

3.36. а. Теперь становится естественной гипотеза, что достаточным условием локального минимума должно быть условие $F_{z'z'}(t, y(t), y'(t))\,k^2 \geq c^2|k|^2$ при любом $t \in T$ и любом $k \in Y$. Это действительно так, но при некотором дополнительном предположении; и является результатом более тонкого анализа, к которому мы и переходим.

б. Функция Якоби-Гамильтона. Рассмотрим в условиях 3.35 пучок Γ экстремалей уравнения Эйлера, выходящий из точки $\{a, y_0\}$. Обозначим эти экстремали через $y = y(t, u)$, где $u \in Y$, $|u - v_0| \leq \varepsilon$ есть угловой коэффициент экстремали при ее выходе из точки $\{a, y_0\}$. При $t = b$ (и при каждом меньшем $t > 0$) эти траектории покрывают взаимно-од-

нозначно некоторую область в пространстве y. Мы не будем предполагать сейчас малости $\beta - \alpha$, а предположим лишь существование такого пучка.

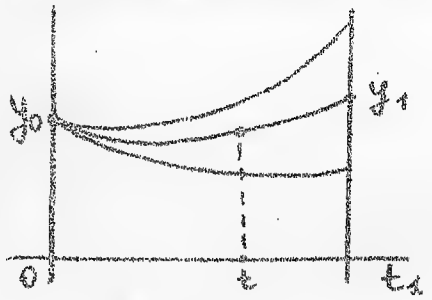

Рис. 3.3-9

В пространстве $T \times y$ траектории пучка покрывают некоторую область G, причем через каждую внутреннюю точку области G проходит одна и только одна траектория этого пучка. Для данной точки $\{t, y\} \in G$ определим функцию

$$S(t,y) = \int_0^t F(\xi, y(\xi), y'(\xi))\, d\xi \qquad (1)$$

где $y(\xi)$ – экстремаль, ведущая в точку $\{t, y\}$. Функция $S(t,y)$ называется функцией Якоби-Гамильтона для функционала $\Phi(y)$. Она зависит от точки y_0, как от параметра.

в. Найдем полный дифференциал функции $S(t,y)$ при переходе из точки $\{t, y\}$ в точку $\{t+\Delta t, y+\Delta y\}$.

Мы имеем при $\Delta t > 0$

$$S(t+\Delta t, y+\Delta y) - S(t,y) = \int_0^{t+\Delta t} F(\xi, y_1(\xi), y_1'(\xi))\, d\xi - \int_0^t F(\xi, y(\xi), y'(\xi))\, d\xi =$$

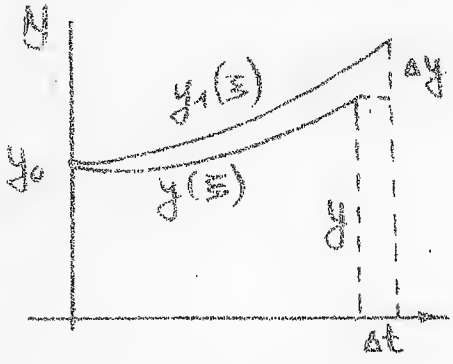

Рис. 3.3-10

$$= \int_0^t [F(\xi, y_1(\xi), y_1'(\xi)) - F(\xi, y(\xi), y'(\xi))]\, d\xi + \int_t^{t+\Delta t} F(\xi, y_1(\xi), y_1'(\xi))\, d\xi, \qquad (2)$$

где $y_1(\xi)$ траектория, ведущая из точки $\{0, y_0\}$ в точку $\{t+\Delta t, y+\Delta y\}$.

Первое слагаемое имеет главную линейную часть, согласно 3.31 (4), равную

$$\int_0^t \left[\frac{\partial F}{\partial y} h(\xi) + \frac{\partial F}{\partial y'} h'(\xi) \right] d\xi \qquad (3)$$

где $h(\xi) = y_1(\xi) - y(\xi)$. Поскольку $y(\xi)$ есть экстремаль, которая, по доказанному, обладает второй производной, в выражении (3) мы можем осуществить интегрирование второго слагае-

мого по частям с переходом от $h'(\xi)$ к $h(\xi)$. Мы получим при этом

$$\int_0^t \left[\frac{\partial \mathcal{F}}{\partial y} h(\xi) + \frac{\partial \mathcal{F}}{\partial z} h'(\xi) \right] d\xi =$$

$$= \int_0^t \left[\frac{\partial \mathcal{F}}{\partial y} - \frac{d}{d\xi} \frac{\partial \mathcal{F}}{\partial z} \right] h(\xi) d\xi +$$

$$+ \frac{\partial \mathcal{F}(\tau, y(\tau), y'(\tau))}{\partial z} h(\tau) =$$

$$= \frac{\partial \mathcal{F}}{\partial z} (\Delta y - y' \Delta\tau + o(\Delta\tau)) \qquad (4)$$

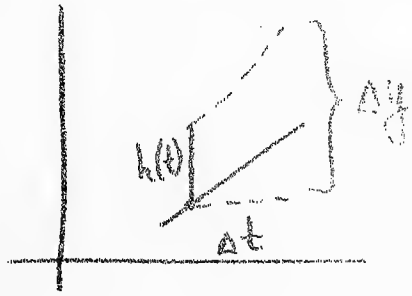

Рис. 3.3-11

поскольку подинтегральное выражение на экстремали $y(\tau)$ обращается в 0.

Второе слагаемое в (2) допускает выделение главной линейной части следующим образом:

$$\int_\tau^{\tau+\Delta\tau} \mathcal{F}(\xi, y+\Delta y, y'+\Delta y') d\xi = \mathcal{F}(\tau, y(\tau), y'(\tau)) \cdot \Delta\tau$$

В итоге можно написать

$$dS(\tau, y) = \mathcal{F}_{z'}(\tau, y, y') dy + (\mathcal{F}(\tau, y, y') - \mathcal{F}_{z'}(\tau, y, y') y') d\tau$$

Следовательно,

$$\frac{\partial S}{\partial y} = \mathcal{F}_{z'}(\tau, y, y') \qquad (5)$$

$$\frac{\partial S}{\partial \tau} = \mathcal{F}(\tau, y, y') - \mathcal{F}_{z'}(\tau, y, y') y' \qquad (6)$$

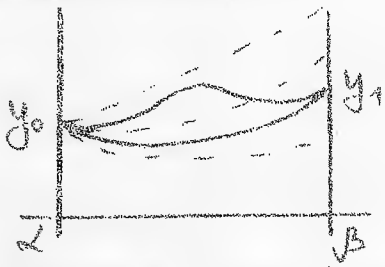

Рис. 3.3-12

причем в правых частях величины y' являются однозначно определенными функциями от τ и y.

г. Рассмотрим теперь какую-либо другую кривую $y_1(\tau)$, также соединяющую точки $\{\alpha, y_0\}$ и $\{\beta, y_1\}$ в пределах области покрытой пучком экстремалей Γ. Мы имеем

$$\phi[y] = S(B, y_1) = \int_{\alpha}^{\beta} F(x, y(x), y'(x))\,dx = \int_{y=y_1(x)} dS =$$

$$= \int_{y=y_1(x)} [F_z(x, \bar{y}(x, y), y'(x))]\,dy_1 + [F(x), \bar{y}(x), y'(x)) - F_z(x, \bar{y}(x), y'(x)) y'(x)]\,dx$$

Здесь аргумент $y'(x)$ означает, в соответствии с определением функции S и вычислением ее дифференциала, угловой коэффициент экстремали, проходящей через точку $\{x, \bar{y}(x)\}$ на кривой γ.

В то же время

$$\phi[\bar{y}] = \int_{\alpha}^{\beta} F(x, \bar{y}(x), \bar{y}'(x))\,dx$$

Таким образом,

$$\phi[\bar{y}] - \phi[y] = \int_{\alpha}^{\beta} [F(x, \bar{y}_1(x), \bar{y}_1'(x)) - F(x, \bar{y}(x), y'(x)) +$$

$$+ F_z(x, \bar{y}(x), y'(x))y'(x) - F_z(x, y(x), y'(x))\bar{y}'(x)]\,dx =$$

$$= \int_{\alpha}^{\beta} \{[F(x, \bar{y}_1(x), \bar{y}_1'(x)) - F(x, \bar{y}(x), y'(x))] - F_z(x, \bar{y}(x), y'(x))[\bar{y}'(x) - y'(x)]\}\,dx$$

Используя разложение Тейлора, равномерное по всем $x \in [\alpha, \beta]$

$$F(x, y, \bar{z}) - F(x, y, z) - F_z(x, y, z)(\bar{z} - z) = \frac{1}{2} F_{zz}(x, y, z)(\bar{z} - z)^2 + o((\bar{z} - z)^2)$$

мы можем записать результат в виде

$$\phi[\bar{y}] - \phi[y] = \frac{1}{2} \int_{\alpha}^{\beta} [F_{zz}(x, \bar{y}(x), y'(x))(\bar{y}'(x) - y'(x))^2 +$$

$$+ o(\sup |\bar{y}'(x) - y'(x)|^2)]\,dx \qquad (7)$$

д. Теперь мы можем сформулировать достаточное условие экстремума:

Теорема. Экстремаль $y = y(x)$, включенная в пучок экстремалей, дает локальный минимум (в пространстве $y^1(T)$) функционалу $\phi[y]$, если форма $F_{zz}(x, y(x), y'(x))\,dh^2 > c\,dh^2$, $c > 0$, при всех $x \in T$.

Доказательство. В силу непрерывности функции $F_{z'z'}(x, y, z)$, существует такое δ, что выполняется при всех x также и неравенство

$$F_{z'z'}(x, y_1(x), y'(x))\,h'h > c\,|h|^2$$

где $|y_1(x) - y(x)| < \delta$. Далее возьмем такое $\varepsilon > 0$, что при $\sup |\bar{y}'(x) - y'(x)| < \varepsilon$ выполняется неравенство

$$o\left(\sup |y_1'(x) - y'(x)|^2\right) < \frac{c}{2}\,|y_1'(x) - y'(x)|^2$$

где $o(\dots)$ есть второе слагаемое под знаком интеграла (7). Тогда, беря $|\bar{y}(x) - y(x)| < \delta$ и $|\bar{y}'(x) - y'(x)| < \varepsilon$, мы получим

$$\Phi[\bar{y}] - \Phi[y] > \frac{1}{2}\int_\alpha^\beta \left[c\,|\bar{y}'(x) - y'(x)|^2 - \frac{c}{2}\,|\bar{y}'(x) - y'(x)|^2\right] dx,$$

$$> \frac{c}{4}\int_\alpha^\beta |\bar{y}'(x) - y'(x)|^2\,dx > 0,$$

т.е. $\Phi[\bar{y}] > \Phi[y]$, если $|\bar{y}(x) - y(x)| < \delta$ и $|\bar{y}'(x) - y'(x)| < \varepsilon$.

Теорема доказана.

Подчеркнем существенность условия "экстремаль, включенная в пучок экстремалей". Оно было весьма существенно при доказательстве, поскольку мы использовали свойства функции Якоби-Гамильтона, определенной в области, покрытой пучком экстремалей. Для малых дуг экстремалей оно всегда выполняется, так как малую дугу экстремали, как мы видели, всегда можно включить в пучок экстремалей; но для длинных дуг оно может не выполняться, и тогда, даже при выполнении условия $F_{z'z'}\,h'h > c\,|h|^2$, данная экстремаль может не давать локального минимума.

3.37. Рассмотрим с точки зрения выполнения достаточных условий примеры, приведенные в 3.33.

а. Функционал 3.33б

$$\Phi[y] = \int_\alpha^\beta \sqrt{1 + y'^2}\,dx.$$

Здесь экстремали - прямые и данная экстремаль всегда включается в пучок экстремалей. Далее

$$F(z) = \sqrt{1+z^2},\quad F_z = \frac{z}{\sqrt{1+z^2}},\quad F_{zz} = \frac{E}{(1+z^2)^{\frac{3}{2}}}$$

представляет собой положительный оператор, $F_{zz} = k^2 > c|k|^2$ в любой ограниченной области изменения z, поэтому достаточное условие действует: отрезок прямой действительно дает локальный (и даже, конечно, абсолютный) минимум длин кривых, соединяющих две дальние точки.

б. Функционал 3.33в

$$\Phi[y] = \int_{x=a}^{b} \frac{\sqrt{1+y'^2}}{v(x)}\,dx$$

Здесь экстремаль также включается в пучок экстремалей. Мы имеем далее

$$F(x,z) = \frac{\sqrt{1+z^2}}{v(x)},\quad F_{zz}(x,z) = \frac{E}{v(x)\,(1+z^2)^{\frac{3}{2}}}$$

эта величина также положительная, следовательно, экстремаль реализует локальный минимум времени.

в. Функционал 3.33г $\Phi[y] = \int_{a}^{a} y\sqrt{1+y'^2}\,dx,\ y(-c) = \rho,$

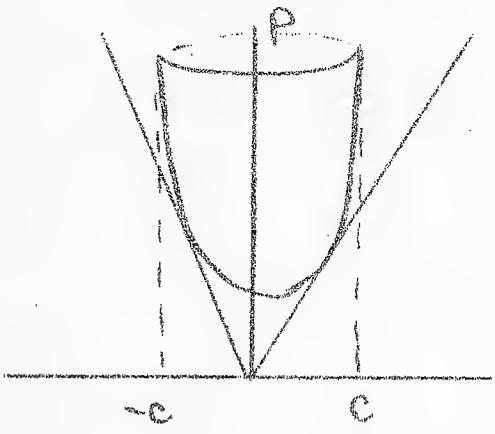

Рис. 3.3-14

Из двух экстремалей, соединяющих точки $(-c,\rho)$ и (c,ρ), верхнюю можно включить в пучок экстремалей, пользуясь сдвигами однопараметрического семейства экстремалей $y = c\,ch\frac{x}{c}$;

Далее мы имеем на экстремали $y(x)$

$$F_{zz} = \frac{y}{(1+z^2)^{\frac{3}{2}}} > c > 0$$

Следовательно, верхняя экстремаль реализует локальный минимум. Нижнюю экстремаль нельзя включить в пучок экстремалей и признак Вейерштрасса не дает ответа. Более точный анализ (см.

Ахиезер, Лекции по вариационному исчислению, 1955, стр.189), показывает, что нижняя экстремаль не дает экстремума.

г. Для уравнений механики, связанных с функционалом 3.33д

$$\int_{\alpha}^{\beta} L\,dt = \int_{\alpha}^{\beta} (T-U)\,dt,$$

где $T = \frac{1}{2}\sum m_i \dot{x}_i^2$, $U = U(x_1, \dots, x_n)$ мы имеем

$$L_{\dot{z}\dot{z}} \cong \begin{Vmatrix} m_1 & & 0 \\ & \ddots & \\ 0 & & m_n \end{Vmatrix}$$

Этот оператор, очевидно, положительный $(L_{\dot{z}\dot{z}}\,k k = \sum m_i k_i^2)$, поэтому для малых экстремалей (которые всегда можно окружить пучком) интеграл действия фактически минимален среди всех близких кривых. Поэтому соответствующий вариационный принцип называют также принципом наименьшего действия. Но следует иметь в виду, что для длинных траекторий интеграл действия на траектории движения может уже не давать локального минимума.

3.38. Уравнения Гамильтона. а. Вернемся к уравнениям для функции Якоби-Гамильтона 3.36 (5)-(6)

$$\left.\begin{aligned} \frac{\partial S}{\partial y} &= F_{\dot{z}}(t, y, y') \\ \frac{\partial S}{\partial t} &= F(t, y, y') - F_{\dot{z}}(t, y, y')\cdot y' \end{aligned}\right\} \quad (I)$$

Правые части этих уравнений обозначим следующим образом:

$$F_{\dot{z}}(t, y, y') \equiv p, \qquad (2)$$

$$F(t, y, y') - F_{\dot{z}}(t, y, y')\,y' \equiv -H \qquad (3)$$

Величина p при фиксированных t и y принадлежит к пространству y', сопряженному к пространству y. Если, как мы всегда предполагаем, оператор $F_{\dot{z}\dot{z}}$ обратим, то из уравнения $p = \frac{\partial F}{\partial \dot{z}}$ можно локально выразить $\dot{z}$ через t, y и p. Функцию H, первоначально заданную, как функцию от t, y и y', можно также считать функцией от t, y и p, если подставить значение $\dot{z}$, как функцию от t, y и p. Мы сохраним запись $H(t, y, p)$ и будем

называть эту функцию функцией Гамильтона.

б. В переменных t, y, p своеобразно записываются и сами уравнения Эйлера. Мы имеем из (3), помня, что при образовании дифференциала безразлично, какие аргументы независимые переменные, какие - функции:

$$dH = -dF + d(pz) = -\frac{\partial F}{\partial t}dt - \frac{\partial F}{\partial y}dy - \frac{\partial F}{\partial z}dz + p\,dz + dp\cdot z = -\frac{\partial F}{\partial t}dt - \frac{\partial F}{\partial y}dy + dp\cdot z$$

Считая H функцией от t, y и p, получаем

$$\frac{\partial H}{\partial t} = -\frac{\partial F}{\partial t},\ \frac{\partial H}{\partial y} = -\frac{\partial F}{\partial y},\ \frac{\partial H}{\partial p} = z. \tag{4}$$

При фиксированных t, y и p величина $\frac{\partial H}{\partial t}$ лежит в R_1, $\frac{\partial H}{\partial y}$ - в Y' и $\frac{\partial H}{\partial p}$ - в Y'', сопряженном к Y'; однако, последнее из уравнений (4) показывает, что на самом деле $\frac{\partial H}{\partial p} = z$ лежит в (меньшем, вообще говоря) пространстве Y.

Уравнение Эйлера имеет вид

$$F_y(t, y(t), y'(t)) = \frac{d}{dt}F_z(t, y(t), y'(t)) \tag{5}$$

Используя соотношения (2) и (4) и заменяя z на $y'(t)$, мы можем теперь заменить (5) двумя уравнениями первого порядка

$$\left.\begin{aligned}\frac{dp(t)}{dt} &= -H_y(t, y(t), p(t))\\ \frac{dy(t)}{dt} &= H_p(t, y(t), p(t))\end{aligned}\right\} \tag{6}$$

Уравнения (6) называются уравнениями Гамильтона, отвечающими функции $F(t, y, z)$

в. Посмотрим, что представляет собою функция H для уравнений механики, где $F = L - U$. При этом мы имеем

$$p = \frac{\partial F}{\partial z} = \frac{\partial T}{\partial z} = \frac{\partial}{\partial z}\left(\frac{1}{2}\sum_{i=1}^{n} m_i z_i^2\right)$$

или $$p_i = m_i z_i,\ z_i = \frac{p_i}{m_i};$$

отсюда

$$H = -F + p\dot{z} = -L + U + \sum \frac{p_i^2}{m_i} = -\frac{1}{2}\sum_{i=1}^{n} m_i \frac{p_i^2}{m_i^2} + U + \sum_{i=1}^{n} \frac{p_i^2}{m_i} =$$

$$= U + \frac{1}{2}\sum_{i=1}^{n} \frac{p_i^2}{m_i} = U + T = E,$$

где E полная энергия системы; функция Гамильтона H в данном случае совпадает с полной энергией системы.

г. Мы видели, что в случае, когда функция Лагранжа L не зависит явно от времени, полная энергия E механической системы является первым интегралом уравнений движения, т.е. постоянна на каждой траектории. Оказывается, что аналогичный факт справедлив для любого функционала 3.31 (2): если функция F не зависит явно от t, то функция Гамильтона H доставляет первый интеграл уравнения Эйлера. Пусть $F(t, y, z)$ не зависит явно от t, т.е. $F(t, y, z) = F(y, z)$. Тогда решение $z = z(p, y)$ уравнения $p = \frac{\partial F}{\partial z}$ также явно не зависит от t; следовательно, и функция $H = -\mathcal{F}(y, z) + pz$ ($z = z(p, y)$) явно не зависит от t. Поэтому

$$\frac{dH}{dt} = \frac{\partial H}{\partial y}\dot{y} + \frac{\partial H}{\partial p}\dot{p} = \frac{\partial H}{\partial y}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial y}\frac{\partial H}{\partial p} = 0$$

В вычитаемом мы переставили $\frac{\partial H}{\partial p} \in Y^*$, $\dot{p} = \frac{\partial H}{\partial y} \in Y'$, чтобы, как обычно, знак функционала стоял перед знаком элемента, на который действует этот функционал. Итак, функция H сохраняет свое значение на каждой траектории уравнения Эйлера, что и утверждалось.

3.39. а. Другие первые интегралы можно получить с помощью набора функций Гамильтона-Якоби. В 3.37 мы рассматривали одну такую функцию; но на самом деле их много - например, меняя y_0, мы уже получаем целое семейство пучков экстремалей и соответственно целое семейство функций Гамильтона-Якоби. В переменных t, y, p уравнения для функций Гамильтона-Якоби можно пере-

писать в форме

$$\frac{\partial S}{\partial y} = p,$$

$$\frac{\partial S}{\partial x} = -H(x, y, p);$$

Подставляя первое уравнение во второе мы получаем уравнение (нелинейное уравнение в частных производных)

$$\frac{\partial S}{\partial x} = -H\left(x, y, \frac{\partial S}{\partial y}\right) \qquad (I)$$

Этому уравнению удовлетворяет любая функция Якоби-Гамильтона $S(x, y)$, где бы ни находилась вершина соответствующего пучка экстремалей.

б. Теперь мы будем называть функцией Гамильтона-Якоби любое решение уравнения Гамильтона-Якоби (I).

Пусть известно решение этого уравнения, $S = S(x, y, \alpha)$ зависящее от произвольного параметра α, причем $\frac{\partial S}{\partial \alpha} \neq 0$ Мы утверждаем, что функция $K(x, y) = \frac{\partial S(x, y, \alpha)}{\partial \alpha}$ является первым интегралом уравнений Эйлера, т.е. она постоянна вдоль каждой траектории. Действительно, используя уравнения Гамильтона 3.38 (6), мы получаем

$$\frac{dK(x, y(x))}{dx} = \frac{\partial K}{\partial x} + \frac{\partial K}{\partial y} y'(x) = \frac{\partial}{\partial \alpha}\left[\frac{\partial S(x, y, \alpha)}{\partial x}\right] +$$

$$+ \frac{\partial}{\partial \alpha}\left[\frac{\partial S(x, y, \alpha)}{\partial y}\right] \frac{\partial H}{\partial p} = -\frac{\partial}{\partial \alpha} H\left(x, y, \frac{\partial S(x, y, \alpha)}{\partial y}\right) +$$

$$+ \frac{\partial}{\partial \alpha} \frac{\partial S(x, y, \alpha)}{\partial y} \frac{\partial H}{\partial p} = -\frac{\partial H}{\partial p} \cdot \frac{\partial^2 S(x, y, \alpha)}{\partial y \partial \alpha} + \frac{\partial S(x, y, \alpha)}{\partial \alpha \partial y} \frac{\partial H}{\partial p} =$$

$$= 0.$$

Таким образом, знание достаточно большого запаса функций Гамильтона-Якоби может облегчить нам нахождение решений уравнений Эйлера.